

# Heights on square of modular curves

Pierre Parent (IMB)

July 1, 2016

## Abstract

We develop a strategy for bounding from above the height of rational points of modular curves with values in number fields, by functions which are polynomial in the level. Our main technical tools come from effective arakelovian descriptions of modular curves and jacobians. We then fulfill this program in the following particular case:

If  $p$  is a not-too-small prime number, let  $X_0(p^2)$  be the classical modular curve of level  $p^2$  over  $\mathbb{Q}$ . Assume Brumer's conjecture on the dimension of winding quotients of  $J_0(p)$ . We prove that there is a function  $b(p) = O(p^{13})$  (depending only on  $p$ ) such that, for any quadratic number field  $K$ , the  $j$ -height of points in  $X_0(p^2)(K)$  is less or equal to  $b(p)$ .

AMS 2000 Mathematics Subject Classification 11G18 (primary), 11G05, 11G16 (secondary).

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Curves, jacobians, their quotients and subvarieties</b>	<b>4</b>
<b>3</b>	<b>Arithmetic Chow group of modular curves</b>	<b>10</b>
<b>4</b>	<b><math>j</math>-height and <math>\Theta</math>-height</b>	<b>18</b>
<b>5</b>	<b>Height of modular curves, and the various <math>W_d</math>'s</b>	<b>24</b>
<b>6</b>	<b>Arithmetic Bézout theorem with cubist metric</b>	<b>29</b>
<b>7</b>	<b>Height bounds for quadratic points on <math>X_0(p^2)</math></b>	<b>38</b>

## 1 Introduction

Let  $N$  be an integer,  $\Gamma_N$  a congruence subgroup of  $\mathrm{SL}_2(\mathbb{Z})$ , and  $X_{\Gamma_N}$  the associate modular curve over some subfield of  $\mathbb{Q}(\mu_N)$  (that we assume from now to be  $\mathbb{Q}$ , to simplify the discussion). The genus  $g_N$  of  $X_{\Gamma_N}$  grows roughly as a polynomial function of  $N$ . So if  $N$  is not too small,  $X_{\Gamma_N}$  has only a finite number of rational points with values in any given number field, by Mordell-Faltings. If one is interested in explicitly determining the set of rational points however, that finiteness fact is of course not sufficient, and a much more desirable control would be provided by upper bounds, for some handy height, on those points. Proving such an “effective Mordell” is known to be an extremely hard problem for arbitrary algebraic curves on number fields<sup>1</sup>.

In the case of modular curves, however, the situation is much better. Indeed, whereas the jacobian of a random algebraic curve should be a somewhat equally random irreducible abelian variety, it is well-known that the jacobian  $J_{\Gamma_N}$  of  $X_{\Gamma_N}$  decomposes up to isogeny into a product of

---

<sup>1</sup>Of course S. Mochizuki has announced a proof of the *abc* conjecture.

quotient abelian varieties defined by Galois orbits of cuspforms for  $\Gamma_N$ . Moreover, in many cases, a nontrivial part of those factors happen to have rank zero over  $\mathbb{Q}$ . Call “winding quotient” the largest quotient  $J_{\Gamma_N, e}$  of  $J_{\Gamma_N}$  with trivial  $\mathbb{Q}$ -rank. Our rustic starting observation is the following: if

$$X_{\Gamma_N} \xrightarrow{\iota} J_{\Gamma_N} \xrightarrow{\pi_e} J_{\Gamma_N, e}$$

is some Albanese map from the curve to its jacobian followed by the projection to  $J_{\Gamma_N, e}$ , then any rational point on  $X_{\Gamma_N}$  has an image which is a torsion (because rational) point on  $J_{\Gamma_N, e}$ , hence has 0 (normalized) height. The pull-back of the invertible sheaf defining the (say) theta height on  $J_{\Gamma_N, e}$  therefore defines a height on  $X_{\Gamma_N}$  which is trivial on rational points. That height in turn necessarily compares to any other natural one, for instance the modular  $j$ -height. Therefore the  $j$ -height of any rational point on  $X_{\Gamma_N}$  is also zero “up to error terms”. Making those error terms explicit would give us the desired upper bound for the height of points on  $X_{\Gamma_N}$ . With some technical complications, that approach can in principle be generalized to degree- $d$  number fields, by considering rational points on symmetric powers  $X_{\Gamma_N}^{(d)}$  of  $X_{\Gamma_N}$  (at least if  $\dim J_{\Gamma_N, e} \geq d$ ). To be a little bit more precise in the present case of symmetric squares, let us associate to a quadratic point  $P$  in  $X_0(p)$  the  $\mathbb{Q}$ -point  $Q := (P, {}^\sigma P)$  of  $X_0(p)^{(2)}$ . Its image  $\iota(Q)$  via some appropriate Albanese embedding in  $J_0(p)$  lies above a torsion point  $a$  in  $J_e$ : assume for simplicity  $a = 0$ . We therefore know  $\iota(Q)$  belongs to the intersection of  $\iota(X_0(p)^{(2)})$  with the kernel  $\tilde{J}_e^\perp$  of the projection  $\pi_e: J_0(p) \twoheadrightarrow J_e$ . To improve the situation, we can even state  $\iota(Q)$  lies in the intersection of  $\iota(X_0(p)^{(2)})$  with the “projection”, in some appropriate sense, of the latter surface on  $\tilde{J}_e^\perp$ . Then one can show that this intersection is 0-dimensional (but here we need to assume Brumer’s conjecture, see below) so that its theta height is controlled, via some arithmetic Bézout theorem, in terms of the degree and height of the two surfaces we intersect. Using an appropriate version of Mumford’s repulsion principle one derives a bound for the height of  $\iota(P)$  too (and not only for its sum  $\iota(Q)$  with its Galois conjugate). Then one makes the translation again from theta height to  $j$ -height on  $X_0(p)$ .

Of course, a nontrivial technical work is necessary to give sense to the straightforward strategy sketched above. The aim of this article is thus to show the possibility of that approach, by making it work in what we feel to be the simplest non-trivial case: that of quadratic points of the classical modular curve  $X_0(p)$  as above (or rather  $X_0(p^2)$ , for technical reasons), with  $p$  a prime number<sup>2</sup>. In the course of the proof we are led to assume the already mentioned conjecture of Brumer, which asserts that the winding quotient of  $J_0(p) := J_{\Gamma_0(p)}$  has dimension roughly half that of  $J_0(p)$ . That conjecture is in line with what is known on the average rank of elliptic curves; it is also implied by other conjectures in analytic number theory; and a lower bound of  $1/4$  (instead of  $1/2$ ) for the asymptotic ratio  $(\dim J_e / \dim J_0(p))$  has been proven by Iwaniec-Sarnak and Kowalski-Michel-VanderKam. (Actually,  $(1/3 + \varepsilon)$  would be sufficient for us.) In any case, at the moment we cannot get rid of this assumption. (Note it can in principle be numerically checked in all specific cases.)

In this setting, our main result is the following (see Theorem 7.5 below).

**Theorem 1.1** *Assume Brumer’s conjecture (cf. Section 2, (18)). Then the quadratic points of  $X_0(p^2)$  have  $j$ -height bounded from above by  $O(p^{13})$ .*

Needless to say, this result cries for both sharpening and generalization, which we intend to tackle in the future. Yet it should be possible to immediately use avatars of Theorem 1.1 to prove the vacuity of sets of rational points of some specific modular curves of arithmetic interest. If combined with lower bounds for heights furnished by isogeny theorems as in [6], the above theorem already has consequences on rational points (Corollary 7.6).

Regarding past works about rational points on modular curves, one can notice that most of them use, at least in parts, some variants of Mazur’s method, which can very roughly be divided

---

<sup>2</sup>Note that Larson and Vaintrob have proven, under the GRH, the asymptotic triviality of rational points on  $X_0(p)$  with values in any given number field which does not contain the Hilbert class field of some quadratic imaginary field (see [34], Corollary 6.5). Independently of any conjecture, Momose had already proven the same result in the case when  $K$  is a given quadratic number field ([42]). Our method however is uniform in the degree and, more importantly, should generalize to other congruence subgroups.

into two steps: first, map modular curves to winding quotients as described above; then prove some quite delicate properties about completions of that map to  $J_e$  (formal immersion criteria). That second step is probably the one which is the most difficult to carry over to great generality. The method we here propose therefore allows to make use of the only first and crucial fact - the mere existence of nontrivial winding quotients. That is known to be, for many cases, a deep result of Kolyvagin-Logachev-Kato, à-la-Birch-Swinnerton-Dyer Conjecture which, again, seems to reflect, from the arithmetic point of view, the quite special properties of the image locus (in the moduli space of principally polarized abelian varieties) of modular curves, among all algebraic curves, under Torelli's map.

In fact the existence of non-trivial winding quotients is not true for all  $\Gamma_N$ , but one can hope to turn that difficulty in the following manner. If  $J_{\Gamma_N}$  has no nontrivial winding quotient over  $\mathbb{Q}$ , one should be able to make a small perturbation of the level structure to force winding quotients to pop-up, then proceed within that broadened setting. More concretely, assume for instance that  $X_{\Gamma_N}$  is the modular curve  $X_{\text{ns}}^+(p)$  over  $\mathbb{Q}$  associate with the normalizer of some non-split Cartan subgroup in prime level  $p$ . Then one knows  $\text{Jac}(X_{\text{ns}}^+(p))$  has precisely no nontrivial winding quotient. However, such a quotient does show-up if one replaces  $X_{\text{ns}}^+(p)$  by  $X_{\text{ns},0}^+(p, 2) := X_{\text{ns}}^+(p) \times_{X(1)} X_0(2)$ , as was first remarked by Darmon and Merel in [16]. One can even hope that the dimension of this winding quotient goes to infinity when  $p$  grows<sup>3</sup>. So starting with a point in  $X_{\text{ns}}^+(p)(\mathbb{Q})$ , say, one can build another one in  $X_{\text{ns},0}^+(p, 2)(K)$ , for  $K$  some cubic number field. Working out our general strategy with  $d = 3$  one can hope to recover our specific “modular effective Mordell” for  $X_{\text{ns}}^+(p)$ .

The methods used in this paper are mainly explicit arakelovian techniques for modular curves and abelian varieties. Such techniques and results have been pioneered, as far as we know, by Abbes, Michel and Ullmo at the end of the 1990s (see in particular [2], [41] and [60], whose results we here eagerly use). They have subsequently been revisited and extended in the work developed by Edixhoven and his school, as mainly (but not exhaustively) presented in the orange book [15]. The tools developed there were motivated by algorithmic Galois-representations issues, but they are very fit for our rational points questions, as shall be clear all through the coming pages. We similarly hope that the effective arakelovian results about modular curves and jacobians we work out here shall prove useful in other contexts<sup>4</sup>.

The layout of this article is as follows. In Section 2 we start gathering classical instrumental facts on quotients of modular jacobians and minimal regular models of  $X_0(p)$  over algebraic integers. In Section 3 we make a precise description of the arithmetic Chow group of  $X_0(p)$ . Section 4 provides an explicit comparison theorem between  $j$ -heights and pull-back of normalized theta height on the Jacobian. Section 5 computes the degree and Faltings height of the image of symmetric products within modular jacobians. In Section 6 we prove our arithmetic Bézout theorem (in the sense of [10]) for general abelian varieties, relative to cubist metrics (instead of the more usual Fubini-Study metrics). This seems more natural, and has the advantage of being quantitatively way more efficient; that is the technical heart of the present paper. Then we apply this arithmetic Bézout to our modular jacobian after technical computations on metric comparisons. Section 7 concludes the computations of the height bounds for quadratic rational points on  $X_0(p^2)$  by making various intersections, projections and manipulations for which to refer to loc. cit.

**Acknowledgments** This work benefited from hours of discussions with Pascal Autissier, who shared with great generosity his expertise in Arakelov geometry, gave us extremely valuable advices, references, explanations, critics, insights, and even red large parts of preliminary releases of the present paper<sup>5</sup>. Many thanks are also due to Qing Liu for clarifying some points of algebraic geometry, Fabien Pazuki for explaining diophantine geometry issues in general and to Gaël

<sup>3</sup>Samuel Le Fourn is currently working on this.

<sup>4</sup>For recent investigations related to more general questions of effective bounds of algebraic points on curves, one can check [13].

<sup>5</sup>... although, as goes without saying, he bears no responsibility for the mistakes which remain.

Rémond, in particular for describing us his own approach to Vojta's method, which under some guise plays a crucial role here.

As already stressed, the influence of the orange book [15] shall be obvious all over this text. We have used many results of the deep effective Arakelovian study of modular curves led there by Bas Edixhoven, Jean-Marc Couveignes and their coauthors. Actually, we also benefited from a visit to Leiden University in June of 2015, where we had very enlightening discussions with Bas, Peter Bruin, Robin de Jong and David Holmes.

Olga Balkanova, Samuel Le Fourn and Guillaume Ricotta helped a lot with references and explanations about some results of analytic number theory, and Jean-Benoît Bost very kindly answered some questions about his own arithmetic Bézout theorem.

**Convention.** In order to avoid numerical troubles, we safely assume in all what follows that primes are by definition strictly larger than 17.

## 2 Curves, jacobians, their quotients and subvarieties

### 2.1 Abelian varieties

#### 2.1.1 Decompositions

Let  $K$  be a field,  $J$  an abelian variety over  $K$  and  $\mathcal{L}$  an ample invertible sheaf defining a polarization of  $J$ . Assume  $J$  is  $K$ -isogenous to a product of two (nonzero) subvarieties, that is, there are (connected...) abelian subvarieties

$$\iota_A: A \hookrightarrow J, \iota_B: B \hookrightarrow J$$

endowed with the polarization  $\iota_A^*(\mathcal{L})$  and  $\iota_B^*(\mathcal{L})$  respectively, such that  $\iota_A + \iota_B: A \times B \rightarrow J$  is an isogeny. Then  $\pi_A: J \rightarrow A' := J \bmod (B)$ , and similarly  $\pi_B: J \rightarrow B'$ , are called *optimal quotients* of  $J$ .

To simplify things, we also assume from now on that  $\text{End}_K(A, B) = \{0\}$ . The product isogeny  $\pi := \pi_A \times \pi_B: J \rightarrow A' \times B'$  induces isogenies  $A \rightarrow A'$  and  $B \rightarrow B'$ . We write

$$\Phi: A \times B \rightarrow J \rightarrow A' \times B'$$

for the obvious composition. Taking (for instance) dual isogenies of  $A \rightarrow A'$  and  $B \rightarrow B'$  we also define an endomorphism

$$\Psi: J \rightarrow A' \times B' \rightarrow A \times B \rightarrow J.$$

When  $K = \mathbb{C}$ , the above constructions are transparent. There is a  $\mathbb{Z}$ -lattice  $\Lambda$  in  $\mathbb{C}^g$ , endowed with a symplectic pairing, such that  $J(\mathbb{C}) \simeq \mathbb{C}^g / \Lambda$  and one can find a direct sum decomposition  $\mathbb{C}^g = \mathbb{C}^{g_A} \oplus \mathbb{C}^{g_B}$  such that if  $\Lambda_A = \Lambda \cap \mathbb{C}^{g_A}$  (and same with  $\Lambda_B$ ) then

$$A(\mathbb{C}) \simeq \mathbb{C}^{g_A} / \Lambda_A \text{ and } B(\mathbb{C}) \simeq \mathbb{C}^{g_B} / \Lambda_B.$$

If  $p_A: \mathbb{C}^g \rightarrow \mathbb{C}^{g_A}$  and  $p_B: \mathbb{C}^g \rightarrow \mathbb{C}^{g_B}$  are the  $\mathbb{C}$ -linear projections relative to that decomposition, then clearly the analytic description of  $\pi_{A, \mathbb{C}}: J(\mathbb{C}) \rightarrow A'(\mathbb{C})$  is

$$z \bmod \Lambda \mapsto z \bmod (\Lambda + \Lambda_B \otimes \mathbb{R}) = p_A(z) \bmod (p_A(\Lambda)).$$

Summing-up, we have lattices inclusions:  $\Lambda_A \subseteq p_A(\Lambda)$ ,  $\Lambda_B \subseteq p_B(\Lambda)$  in  $\mathbb{C}^g$  such that

$$\Lambda_A \oplus \Lambda_B \subseteq \Lambda \subseteq p_A(\Lambda) \oplus p_B(\Lambda),$$

which induces our isogenies. The lattices  $\Lambda_A \oplus \Lambda_B$  and  $p_A(\Lambda) \oplus p_B(\Lambda)$  are dual to each other (or *perpendicular* to each other, in the sense of [47], p. 126) with respect to the symplectic form  $\mathbf{E}$  on  $\Lambda$ . Recall indeed that, for any subset  $V \subseteq \mathbb{C}^g$ , one defines its dual as

$$V^\wedge := \{w \in \mathbb{C}^g, \forall v \in V, \mathbf{E}(v, w) \in \mathbb{Z}\}.$$

Then one easily checks that

$$(\Lambda_A \oplus \Lambda_B)^\wedge = (p_A(\Lambda) \oplus p_B(\Lambda))$$

(and vice-versa:  $(\Lambda_A \oplus \Lambda_B) = (p_A(\Lambda) \oplus p_B(\Lambda))^\wedge$ ); one even sees that

$$(\Lambda_A)^\wedge = p_A(\Lambda) \oplus (p_B(\Lambda) \otimes \mathbb{R})$$

so that  $\Lambda_A$  and  $p_A(\Lambda)$  are dual lattices of  $T_0(A) \simeq \mathbb{C}^{g_A}$  with respect to the restriction  $\mathbf{E}|_A$  (and same with  $B$ ).

The isogeny  $I'_A: A \rightarrow A'$ , deduced from the inclusion  $\Lambda_A \subseteq p_A(\Lambda)$ , has degree  $\text{card}(p_A(\Lambda)/\Lambda_A)$ . If  $N_A$  is a multiple of the exponent of the quotient  $p_A(\Lambda)/\Lambda_A$ , there is an isogeny  $I_{A,N_A}: A' \rightarrow A$  such that  $I_{A,N_A} \circ I'_A$  and  $I'_A \circ I_{A,N_A}$  both are multiplication by  $N_A$ . The analytic descriptions of the above clearly are:

$$\left\{ \begin{array}{ccc} A(\mathbb{C}) \simeq \mathbb{C}^{g_A}/\Lambda_A & \xrightarrow{I'_A} & A'(\mathbb{C}) \simeq \mathbb{C}^{g_A}/p_A(\Lambda) \\ z & \mapsto & z \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{ccc} \mathbb{C}^{g_A}/p_A(\Lambda) & \xrightarrow{I_{A,N_A}} & \mathbb{C}^{g_A}/\Lambda_A \\ z & \mapsto & N_A z. \end{array} \right. \quad (1)$$

**Remark 2.1** Consider now the case of one only given immersion  $A \hookrightarrow J$ , over a number field. One might apply [23], Théorème 1.3, to deduce the existence of an abelian variety  $B$  such that the degree of  $A \times B \xrightarrow{+} J$ :

$$|A \cap B| = \Lambda/\Lambda_A \oplus \Lambda_B$$

(with our previous notations) is bounded from above by an explicit function  $\kappa(J)$  of the stable Faltings' height  $h_F(J)$ . Note that when  $A$  and  $(J \bmod A)$  are not isogenous (which will be the case for us), then there is actually no choice for that  $B \hookrightarrow J$ : it has to be the Poincaré complement to  $A$ . The isogeny  $J \rightarrow A' \times B'$  given by the two projections has degree  $p_A(\Lambda) \oplus p_B(\Lambda)/\Lambda$ , which also is  $|A \cap B| := N$ . One can therefore take the  $N_A$  intervening just before the present Remark 2.1 as equal to  $N$ , and

$$N \leq \kappa(J).$$

Making the same for  $B' \rightarrow B$ , the preceding morphism  $\Psi$  is then simply the multiplication  $J \xrightarrow{[N]} J$  by the integer  $N$ . Although we will not need numerical estimates for those quantities in what follows, it is not hard, using results by Ullmo et al., to make them explicit in our setting of modular curves and jacobians over number fields.

### 2.1.2 Polarizations and heights

Keeping the above notations and hypothesis, let now  $\Theta$  be an ample sheaf on  $J$ , and  $I_A := I_{A,N}: A' \rightarrow A$  (respectively,  $I'_{B,N}$ ) as in Remark 2.1. We consider the composed morphism:

$$\varphi_A: J \xrightarrow{\pi_A} A' \xrightarrow{I_A} A \xrightarrow{\iota_A} J \quad (2)$$

and pull-back the sheaf  $\Theta$  along it: the immersion  $\iota_A: A \hookrightarrow J$  defines a polarization  $\Theta_A := \iota_A^*(\Theta)$  on  $A$ , whence a polarization  $\Theta_{A'} := I_A^*(\Theta_A)$  on  $A'$ , and finally an invertible sheaf  $\Theta_{J,A} := \pi_A^*(\Theta_{A'})$  on  $J$ . Now composing the morphisms:

$$J \xrightarrow{\pi_A \times \pi_B} A' \times B' \xrightarrow{I_A \times I_B} A \times B \xrightarrow{\iota_A + \iota_B} J \quad (3)$$

gives the multiplication-by- $N$ :  $J \xrightarrow{[N]} J$ . Therefore, assuming for simplicity  $\Theta$  is symmetric, one has

$$[N]^*\Theta \simeq \Theta^{\otimes N^2} \simeq \Theta_{J,A} \otimes_{\mathcal{O}_J} \Theta_{J,B}. \quad (4)$$

If  $K$  is a number field, the Néron-Tate normalization process associates with  $\Theta$  a system of compatible euclidean norms  $h_\Theta = \|\cdot\|_\Theta^2$  on the finite-dimensional  $\mathbb{Q}$ -vector spaces  $J(F) \otimes_{\mathbb{Z}} \mathbb{Q}$ , for  $F/K$  running through the number field extensions of  $K$ , and similarly euclidean norms  $h_{\Theta_A} :=$

$\|\cdot\|_{\Theta_A \otimes \frac{1}{N^2}}^2 := \frac{1}{N^2} \|\cdot\|_{\Theta_A}^2$  on  $A(F) \otimes_{\mathbb{Z}} \mathbb{Q}$  and  $h_{\Theta_B} := \frac{1}{N^2} \|\cdot\|_{\Theta_B}^2$  on  $B(F) \otimes_{\mathbb{Z}} \mathbb{Q}$  such that, under the isomorphisms  $J(F) \otimes_{\mathbb{Z}} \mathbb{Q} \simeq (A(F) \otimes_{\mathbb{Z}} \mathbb{Q}) \oplus (B(F) \otimes_{\mathbb{Z}} \mathbb{Q})$ , one has

$$h_{\Theta} = h_{\Theta_A} + h_{\Theta_B}. \quad (5)$$

We will later on need to interpret the above projections of type  $J \xrightarrow{\pi_A} A'$  from (2) as sort of internal operations. This is of course not possible as such, but we shall use the following artefact. Recall from (1) the definition of  $N_A$ , that of the maps  $A' \xrightarrow{I_{A,N_A}} A$  and  $A \xrightarrow{\iota_A} J$ . Denote by  $[N_A]_A$  the multiplication by  $N_A$  restricted to  $A$ . If  $V$  is a closed algebraic subvariety of  $J$ , define

$$\mathcal{P}_A(V) := (\iota_A [N_A]_A^{-1} I_{A,N_A} \pi_A)(V) \quad (6)$$

as the reduced closed subscheme with relevant support. This would therefore just be the projection of  $V$  on  $A$  if  $J$  were *isomorphic* to the product  $A \times B$ , and is the best approximation to that projection in our case when the decomposition of the abelian variety is true only up to isogeny.

Note that  $\mathcal{P}_A(V)$  is a priori highly non-connected. However, all its irreducible geometric components are obtained from each other by translation by a  $N_A$ -torsion point of  $A(\overline{\mathbb{Q}})$ , and for our later purpose we will have the possibility to replace  $\mathcal{P}_A(V)$  by one of its components containing a specific point, say  $P_0$ : we shall denote that component by  $\mathcal{P}_A(V)_{P_0}$ .

Suppose now  $J \sim A \times B$  as above is the jacobian of an algebraic curve  $X$  on  $K$  with positive genus  $g$ . For  $P_0$  a point of  $X(K)$  (or more generally a  $K$ -divisor of degree 1 on  $X$ ) let

$$\iota_{P_0}: \begin{cases} X & \hookrightarrow J \\ P & \mapsto (P) - (P_0) \end{cases} \quad (7)$$

be the Albanese embedding associated with  $P_0$ . We define the classical Theta divisor  $\theta$  on  $J$  which is the image of  $\iota_{P_0}^{g-1}: X^{g-1} \rightarrow J$  and its symmetric version

$$\Theta := (\theta \otimes_{\mathcal{O}_J} [-1]^* \theta)^{\otimes \frac{1}{2}} \quad (8)$$

(which is a translate of  $\theta$  obtained as  $\iota_{\kappa}(X^{g-1})$ , where  $\iota_{\kappa} = t_{\kappa}^* \iota_{P_0}$  for  $t_{\kappa}$  the translation by some  $\frac{\kappa}{2g-2}$  where  $\kappa$  is the canonical divisor on  $X$ . Note that  $\Theta$  does not need to be defined over  $\mathbb{Q}$ ). Our first aim will be to compare, when  $X$  is a modular curve, the height functions  $\|\iota_{P_0}(\cdot)\|_{\Theta_A \otimes \frac{1}{N^2}}$  on  $X(F)$  with another natural height given by the modular  $j$ -function.

We will discuss below an Arakelovian description of Néron-Tate height (see Section 3). We conclude this paragraph by a few remarks as a preparation. Let  $\mathcal{B} := \{\omega_1, \dots, \omega_g\}$  be a basis of  $H^0(X(\mathbb{C}), \Omega_{X/\mathbb{C}}^1) \simeq H^0(J(\mathbb{C}), \Omega_{J/\mathbb{C}}^1)$ , which is orthogonal with respect to the norm

$$\|\omega\|^2 = \frac{i}{2} \int_{X(\mathbb{C})} \omega \wedge \overline{\omega}.$$

The transcendent writing-up of the Abel-Jacobi map  $\iota_{P_0}: P \mapsto (\int_{P_0}^P \omega_i)_{1 \leq i \leq g}$  shows that the pull-back to  $X(\mathbb{C})$  of the translation-invariant measure on  $J(\mathbb{C})$ , normalized to have total mass 1, is

$$\mu_0 = \frac{i}{2g} \sum_{\mathcal{B}} \frac{\omega \wedge \overline{\omega}}{\|\omega\|^2}. \quad (9)$$

More generally,  $\pi_A \circ \iota_{P_0}$  is, over  $\mathbb{C}$ , the map  $P \mapsto (\int_{P_0}^P \omega)_{\omega \in \Omega_A}$ , where  $\mathcal{B}_A$  is some basis of  $H^0(A'(\mathbb{C}), \Omega_{A'/\mathbb{C}}^1) \simeq H^0(J(\mathbb{C}), \pi_A^*(\Omega_{A'/\mathbb{C}}^1)) \subseteq H^0(J(\mathbb{C}), \Omega_{J/\mathbb{C}}^1)$ . Therefore, writing  $g_A := \dim(A') = \dim(A)$ , the pull-back to  $X(\mathbb{C})$  of the translation-invariant measure on  $A'(\mathbb{C})$  (normalized so to have total mass 1 on the curve again) is

$$\mu_A = \frac{i}{2g_A} \sum_{\mathcal{B}_A} \frac{\omega \wedge \overline{\omega}}{\|\omega\|^2}. \quad (10)$$

## 2.2 Modular curves

Here we recall a few classical facts on the minimal regular model of the modular curve  $X_0(p)$ , for  $p$  a prime number, over a ring of algebraic integers. The first general reference on this topic is [17]; see also [15] or [38], [39] and references therein.

### 2.2.1 The $j$ -height

The quotient of the completed Poincaré upper half-plane  $\mathcal{H} \cup \mathbb{P}^1(\mathbb{Q})$  by the classical congruence subgroup  $\Gamma_0(p)$  defines a Riemann surface  $X_0(p)(\mathbb{C})$  which is known to have a geometrically connected smooth and proper model over  $\mathbb{Q}$ . All through this paper, we denote its genus by  $g$ .

The first technical theme of this article is the explicit comparison of various heights on  $X_0(p)(\overline{\mathbb{Q}})$ . When  $V$  is an algebraic variety over a number field  $K$ , any finite  $K$ -map  $\varphi: V \rightarrow \mathbb{P}_K^N$  to some projective space defines a naive Weil height on  $V(\overline{K})$ . This applies in particular when  $V$  is a curve and  $\varphi$  is the finite morphism defined by an element of the function field of  $V$ , and in the case of a modular curve  $X_\Gamma$  associated with some congruence subgroup  $\Gamma$ , say, a natural height to choose on  $X_\Gamma(\overline{\mathbb{Q}})$  is precisely Weil's height  $h(P) = h(j(P))$  relative to the classical  $j$ -function. The degree of the associate map  $X_\Gamma \rightarrow X(1) \simeq \mathbb{P}^1$  is  $[\mathrm{PSL}_2(\mathbb{Z}) : \Gamma]$ , so that number is the class of our Weil height in the Néron-severi group  $\mathrm{NS}(X_\Gamma)$  identified with  $\mathbb{Z}$ . More explicitly if  $X = X_\Gamma$  is defined over the number field  $K$ , say, the  $j$ -morphism is

$$\begin{array}{ccccccc} \begin{array}{c} X \\ P \end{array} & \begin{array}{c} \xrightarrow{j} \\ \mapsto \end{array} & \begin{array}{c} \mathbb{P}_K^1 = \mathrm{Proj}(K[X_0, X_1]) \\ (1, j(P)) = (1/j(P), 1) \end{array} & = & \begin{array}{c} \overline{\mathbb{A}_K^1}^{\mathrm{Zar}} \\ j(P) \end{array} & = & \begin{array}{c} \overline{\mathrm{Spec}}^{\mathrm{Zar}}(K[X_1/X_0]) \\ \frac{X_1}{X_0}(P) \end{array} \end{array}$$

and the Weil height of a point  $P \in X(K)$  is the naive height of its  $j$ -invariant as an algebraic number:

$$h(P) = h(j(P)) = \frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} [K_v : \mathbb{Q}_v] \log(\max(1, |j(P)|_v))$$

which is also Weil's projective height  $h(J(P))$  with respect to the above basis  $(X_0, X_1 = X_0 j)$  of global sections of  $\mathcal{O}_{\mathbb{P}_K^1}(1)$ . Our Weil height on  $X$  is associated with the linear equivalence classes of divisors  $D$  associated with  $J^*(\mathcal{O}_{\mathbb{P}_K^1}(1))$ , so that

$$D \sim (\text{poles of } j \text{ on } X) \sim (\text{zeroes of } j) \sim \sum_{c \in \{\text{cusps of } X\}} e_c \cdot c$$

where each  $e_c$  is the ramification index of  $c$  via  $J$ .

Those considerations lead to explicit comparisons with other heights. Indeed, a more intrinsic way to define heights on algebraic varieties is provided by Arakelov's theory. Defining this properly in the case of our modular curves demands a precise description of regular models for them, which we now recall.

### 2.2.2 Minimal regular models

The normalization of the  $j$ -map  $X_0(p) \rightarrow X(1)_{/\mathbb{Z}} \simeq \mathbb{P}_{/\mathbb{Z}}^1$  over  $\mathbb{Z}$  defines a model for  $X_0(p)$ , that we call the modular model, it is smooth over  $\mathbb{Z}[1/p]$ .

We fix a number field  $K$ , write  $\mathcal{O}_K$  for its ring of integers, and deduce by base change a model for  $X_0(p)$  over  $\mathcal{O}_K$ . After making a few blowing-ups if necessary we obtain a (minimal) regular model of  $X_0(p)$  over  $\mathcal{O}_K$ , which we denote from now on by  $\mathcal{X}_0(p)_{/\mathcal{O}_K}$ , or simply  $\mathcal{X}_0(p)$  if the context prevents confusion. (Just recall here that for  $F/K$  a field extension,  $\mathcal{X}_0(p)_{/\mathcal{O}_F}$  is *not* the base change to  $\mathcal{O}_F$  of  $\mathcal{X}_0(p)_{/\mathcal{O}_K}$  if  $F/K$  ramifies above  $p$ .) Let  $v$  be a place of  $\mathcal{O}_K$  above  $p$ , with residue field  $k(v)$ . The dual graph of  $\mathcal{X}_0(p)$  at  $v$  is well-known to be made of two extremal vertices, which we label  $C_0$  and  $C_\infty$ , containing the cusps 0 and  $\infty$  respectively (see Figure 1). Those two are linked by

$$s := g + 1$$

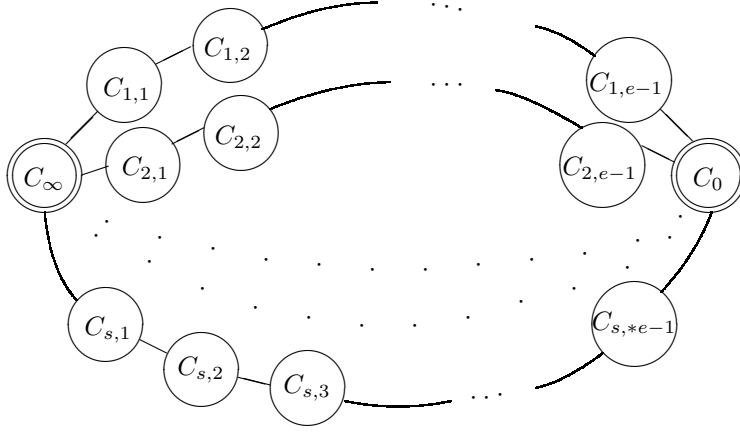


Figure 1: Dual graph of  $\mathcal{X}_0(p)/\mathcal{O}_K$  at  $v$ .

branches. Each branch corresponds to a singular point  $S$  in  $\mathcal{X}_0(p)(\mathbb{F}_{p^2})$ , which in turn parameterizes an isomorphism class of supersingular elliptic curve  $E_S$  in characteristic  $p$ .

The Fricke involution  $w_p$  acts on the dual graph as the continuous isomorphism which exchanges  $C_0$  and  $C_{\infty}$  and acts on the branches as a generator of  $\text{Gal}(\mathbb{F}_{p^2}/\mathbb{F}_p)$ .

We write  $S(1), \dots, S(s)$  for the supersingular points and set

$$w_n := \#\text{Aut}(S(n))/\langle \pm 1 \rangle := \#\text{Aut}_{\mathbb{F}_{p^2}}(E_{S(n)})/\langle \pm 1 \rangle$$

which is equal to 1 except in the (at most two) cases when the underlying supersingular elliptic curve has  $j$ -invariant 1728 or 0, where it is equal to 2 or 3 respectively. Now each path, or branch, on our dual graph at  $v$  passes through  $(w_n e - 1)$  vertices (for  $e$  the ramification index of  $K$  at  $v$ ), that is, again, equal to  $e - 1$  except for at most two branches: one of length  $2e - 1$  (obtained by blowing-up the supersingular point of moduli  $j = 1728 \bmod v$ , if showing-up), and a path of length  $3e - 1$  (obtained by blowing-up, if needed, the supersingular point of moduli  $j = 0 \bmod v$ ). We enumerate the vertices  $\{C_{n,m}\}_{1 \leq m \leq w_n e - 1}$  in the “generic” path  $n$ , and  $\{C_{\{4\},m}\}_{1 \leq m \leq 2e - 1}$ ,  $\{C_{\{3\},m}\}_{1 \leq m \leq 3e - 1}$  for the potential two exceptional paths. We also denote the familiar quantity  $\sum \frac{1}{w_n}$  by  $w(\text{Eis})$ , the sum being taken over the set of all supersingular points of  $\mathcal{X}_0(p)/\mathcal{O}_{K,v}$ . The well-known Eichler mass formula now writes

$$w(\text{Eis}) = \sum_{1 \leq n \leq s} \frac{1}{w_n} = \frac{p-1}{12} \quad (11)$$

(see for instance [26], p. 117). Recall this implies that the genus  $g$  of  $X_0(p)$  is equivalent to  $p/12$  (the exact formula depending on the residue class of  $p \bmod 12$ ) and in any case:

$$\frac{p-13}{12} \leq g \leq \frac{p+1}{12} \quad (12)$$

(see for instance p. 117 of [26] again).

Abusing a bit notations,  $C_{\infty}$  will sometimes be also denoted as  $C_{n,0}$ , and similarly  $C_0$  might be written as  $C_{n,w_n e}$ . We choose as a basis for  $\oplus_{C \in \mathcal{C}} \mathbb{Z} \cdot C$  the ordered set

$$\mathcal{B} = (C_{\infty}, (C_{1,1}, C_{1,2}, \dots, C_{1,e-1}), (C_{2,1}, \dots, C_{2,e-1}), \dots, (C_{s,1}, \dots, C_{s,w_s e-1}), C_0) \quad (13)$$

(that is, we list the vertices by running through each branch successively, and put the possible branches of length twice or thrice the normal length at the end). At bad places  $v$  the intersection matrix restricted to each submodule  $\oplus_{m=1}^{w_n e-1} \mathbb{Z} \cdot C_{n,m}$  (for some fixed branch of index  $n$ ) is then clearly  $(\log(\#k(v))) \cdot \mathcal{M}_0$ , where

$$\mathcal{M}_0 = \begin{pmatrix} -2 & 1 & 0 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & 1 & -2 & 1 \\ 0 & 0 & \cdots & 0 & 1 & -2 \end{pmatrix}, \quad (14)$$



whose only dependence on  $n$  is that its type is  $(w_n e - 1) \times (w_n e - 1)$ . That matrix has determinant  $(-1)^{w_n e - 1} w_n e$ . Define the row vectors:

$$L := (1 \ 0 \ 0 \ \cdots \ 0), \ L' := (0 \ 0 \ 0 \ \cdots \ 1)$$

and the transpose column vectors:

$$V := L^t, \ V' := L'^t.$$

The intersection matrix on the whole space  $\mathbb{Z}^{\mathcal{B}}$  is therefore  $(\log(\#k(v)) \cdot \mathcal{M})$  for

$$\mathcal{M} = \begin{pmatrix} -s & L & L & \cdots & L & 0 \\ V & \mathcal{M}_0 & 0 & \cdots & 0 & V' \\ V & 0 & \mathcal{M}_0 & \cdots & 0 & V' \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ V & 0 & 0 & \cdots & \mathcal{M}_0 & V' \\ 0 & L' & L' & \cdots & L' & -s \end{pmatrix}. \quad (15)$$

(This of course has to be modified in the obvious way when  $e_v = 1$ .)

### 2.2.3 Winding quotients, their dimension

The jacobian of  $X_0(p)_{\mathbb{Q}}$  is usually denoted by  $J_0(p)$ . As  $\mathcal{X}_0(p)$  is semistable over  $\mathbb{Z}$ , the Néron model  $\mathcal{J}_0(p)$  of  $J_0(p)$  is a semi-abelian scheme over  $\mathbb{Z}$  (and an abelian scheme over  $\mathbb{Z}[1/p]$ ) whose neutral components represents the neutral component  $\text{Pic}_{\mathbb{Z}}^0(\mathcal{X}_0(p))$  of the relative Picard functor of  $\mathcal{X}_0(p)$  over  $\mathbb{Z}$ .

We know from Shimura's theory that the natural decomposition of cotangent spaces into Hecke eigenspaces induces a corresponding decomposition over  $\mathbb{Q}$  of abelian varieties up to isogenies:

$$J_0(p) \sim \prod_{f \in S_2^{\text{new}}(\Gamma_0(p))/G_{\mathbb{Q}}} J_f \quad (16)$$

indexed by Galois orbits of weight-2 newforms for  $\Gamma_0(p)$ . A first useful partition of this decomposition comes from the sign of the functional equations for the  $L$ -functions of eigenforms  $f$ , that is, whether  $w_p(f)$  equals  $f$  or  $-f$ . One therefore write  $J_0(p)^-$  for the optimal quotient abelian variety associated with  $\prod_{f, w_p(f) = -f} J_f$  in (16), and similarly  $J_0(p)^+$ . One has  $J_0(p)^- = J_0(p)/(1 + w_p)J_0(p)$  and  $J_0(p)^+ = J_0(p)/(1 - w_p)J_0(p)$ . One knows that

$$\dim J_0(p)^- = \left(\frac{1}{2} + o(1)\right) \dim J_0(p)$$

(cf. e.g. [57], Lemme 3.2).

A more subtle object is the *winding quotient*  $J_e$ , defined as the optimal quotient of  $J_0(p)$  corresponding to  $\prod_{f, L(f, 1) \neq 0} J_f$  in decomposition (16). One can write

$$J_e = J_0(p)/I_e J_0(p) \quad (17)$$

for some ideal  $I_e$  of the Hecke algebra  $\mathbb{T}_{\Gamma_0(p)}$ . Similarly,  $J_e^{\perp}$  will denote the optimal quotient corresponding to  $\prod_{f, L(f, 1) = 0} J_f$ . For obvious reasons regarding signs of functional equations,  $J_e$  is contained in  $J_0(p)^-$ . But more is expected: in line with the principle that “*the vanishing order of a (modular)  $L$  functions at the critical point should generically be as small as allowed by parity*”, Brumer ([12]) conjectured that, as  $p$  tends to infinity,

$$(?) \quad \dim J_e = (1 - o(1)) \dim J_0(p)^-. \quad (\text{Brumer}) \quad (18)$$

Equivalently, it is conjectured that  $\dim J_e = (\frac{1}{2} + o(1)) \dim J_0(p)$ , or that the dimensions of  $J_e$  and  $J_e^{\perp}$  should be, asymptotically in  $p$ , of equal size. Note that (18) above is also implied by the

“Density Conjecture” of [29], p. 56 et seq., cf. loc. cit. Remark F on p. 65<sup>6</sup>. Actually, what we eventually need in this article (cf. Section 7) is a weaker form of (18), which is:

$$(?) \quad \dim J_e \geq \frac{\dim J_0(p)}{3} + 1 \quad \text{for large enough } p. \quad (19)$$

An important theorem of Iwaniec-Sarnak and Kowalski-Michel asserts something quite as good, that is :

$$\left(\frac{1}{4} - o(1)\right) \dim J_0(p) \leq \dim J_e \leq \left(\frac{1}{2} + o(1)\right) \dim J_0(p) \quad (20)$$

as  $p$  goes to infinity (so that  $(\frac{1}{2} - o(1)) \dim J_0(p) \leq \dim J_e^\perp \leq (\frac{3}{4} + o(1)) \dim J_0(p)$ ), see [30], Corollary 13 and [33]). One knows that breaking that  $\frac{1}{4}$  is closely linked to the Landau-Siegel zero problem. Assuming the GRH for  $L$ -functions of modular forms, Iwaniec, Luo and Sarnak prove one can improve  $\frac{1}{4}$  to  $\frac{9}{32}$  ([29], Corollary 1.6, (1.54)).

We conclude that paragraph by recalling an endomorphism-view on winding quotients. A theorem of Ribet ([56]) asserts that the Hecke algebra  $\mathbb{T}_{\Gamma_0(p)} \otimes_{\mathbb{Z}} \mathbb{Q}$  is equal to the full ring of endomorphisms (up to isogeny)  $\text{End}_{\mathbb{Q}}(J_0(p)) \otimes_{\mathbb{Z}} \mathbb{Q}$ . In that case, writing  $A := \tilde{J}_e$ ,  $A' := J_e$ ,  $B := \tilde{J}_e^\perp$  and  $B' := J_e^\perp$  in the setting of (3), there are elements  $t_e$  and  $t_e^\perp$  in  $\mathbb{T}_{\Gamma_0(p)}$  such that (sticking to notations as in loc. cit.) the morphism  $\iota_A \circ I_A \circ \pi_A$  (and respectively in  $B$ ) is multiplication by  $t_e$  (respectively,  $t_e^\perp$ ). So (3) can be written as

$$[\cdot N] = [t_e + t_e^\perp]: J_0(p) \xrightarrow{\pi_{J_e} \times \pi_{J_e^\perp}} J_e \times J_e^\perp \xrightarrow{I_{J_e, N} \times I_{J_e^\perp, N}} \tilde{J}_e \times \tilde{J}_e^\perp \xrightarrow{\iota_{J_e} + \iota_{J_e^\perp}} J_0(p).$$

The central object of this paper will eventually be the maps

$$X_0(p)^{(d)} \rightarrow J_0(p) \rightarrow J_e$$

from symmetric products of  $X_0(p)$  (mainly the curve itself and its square) to the winding quotient.

### 3 Arithmetic Chow group of modular curves

We now give a description of the Arakelov geometry of  $X_0(p)$ , relying on the work of many people: that topic has been pioneered by Abbes, Ullmo and Michel ([2], [41], [60]) and notably developed by Edixhoven, Couveignes and their coauthors (see in particular [15]). We shall also use the work of Bruin ([11]), Jürgen-Kramer ([31]) and Menares ([38], [39]) among others. We refer to those articles and their bibliography for general facts on Arakelov theory (see in particular [14] and [18]).

Let  $\mathcal{X}$  be any regular and proper arithmetic surface over the integer ring  $\mathcal{O}_K$  of a number field  $K$ . Fixing in general smooth hermitian metrics  $\mu$  on the base changes of  $\mathcal{X}$  to  $\mathbb{C}$ , it follows from the basics of Arakelov theory that for any horizontal divisor  $D$  on  $\mathcal{X}$  over  $\mathcal{O}_K$  there are Green function  $g_{\mu, D}$  on each Archimedean completion  $\mathcal{X}(\mathbb{C})$  satisfying the differential equation

$$\Delta g_{\mu, D} = -\delta_D + \deg(D)\mu$$

for  $\Delta = \frac{1}{i\pi} \partial \bar{\partial}$  the Laplace operator and  $\delta_D$  the Dirac distribution relative to  $D_{\mathbb{C}}$  on  $\mathcal{X}(\mathbb{C})$ . The function  $g_{\mu, D}$  is integrable on the compact Riemann surface  $\mathcal{X}(\mathbb{C})$  endowed with its measure  $\mu$ , and uniquely determined up to an additive constant which can be fixed by imposing the normalizing condition that

$$\int_{\mathcal{X}(\mathbb{C})} g_{\mu, D} \mu = 0. \quad (21)$$

<sup>6</sup>Quoting Olga Balkanova (private communication), “Theorem 1.1 in [29] is proved for the test function  $\phi$ , whose Fourier transform is supported on the interval  $[-2, 2]$ . The density conjecture claims that the same results are true without restriction on Fourier transform of  $\phi$ , see formula 1.9 [of loc. cit.]”

When the horizontal divisor  $D$  is a section  $P_0$  in  $\mathcal{X}(\mathcal{O}_K)$ , one will sometimes also use the notation  $g_\mu(P_0, z)$  for  $g_{\mu, P_0}(z)$ . The Green functions relative to fixed smooth  $(1, 1)$ -forms  $\mu$  allow to define an Arakelov intersection product relative to the  $\mu$ , which will be denoted by  $[\cdot, \cdot]_\mu$ , or  $[\cdot, \cdot]$  if there is no ambiguity about the implicit form. (In particular, the index will often be dropped for divisors intersections of which one at least is vertical, where the choice of  $\mu$  does not intervene).

Let  $\mu_0$  be the canonical Arakelov  $(1, 1)$ -form on the Riemann surface  $\mathcal{X}(\mathbb{C})$  (assumed to have positive genus), inducing the “flat metric”. It also corresponds to the pullback, by any Albanese morphism  $\mathcal{X}(\mathbb{C}) \rightarrow \text{Jac}(\mathcal{X}_K)(\mathbb{C})$ , of the “cubist” metric in the sense of Moret-Bailly ([43], more about this shortly) on the jacobian  $\text{Jac}(\mathcal{X}_K)$ , associated with the normalized height  $h_\Theta$ .

We now specialize to the case of  $\mathcal{X}_0(p)$  as in Section 2.2. If  $f$  is a modular form of weight 2 for  $\Gamma_0(p)$ , let  $\|f\|^2$  be its Petersson norm. As in (9) (because newforms are orthogonal in prime level) we have

$$\mu_0 := \frac{i}{2 \dim(J)} \sum_{f \in S_2(\Gamma_0(p))^{\text{new}}} \frac{f \frac{dq}{q} \wedge \overline{f \frac{dq}{q}}}{\|f\|^2}. \quad (22)$$

For later use we also consider the Néron-Tate height  $h_{\Theta_e}$  on  $J_e$  (as in (5) and around, for  $A' = J_e$ ) which induces a height  $h_{\Theta_e} \circ \iota_{e, P_0}$  on  $X_0(p)$  via the map  $\iota_{e, P_0}: X_0(p) \rightarrow J \rightarrow J_e$ . The curvature form of the hermitian sheaf on  $X_0(p)$  defining the Arakelov height associated with  $h_{\Theta_e} \circ \iota_{e, P_0}$  is  $\frac{i}{2} \sum_{S[I_e]} \frac{f \frac{dq}{q} \wedge \overline{f \frac{dq}{q}}}{\|f\|^2}$ , where  $S[I_e] = S_2(\Gamma_0(p))^{\text{new}}[I_e]$  is the set of newforms killed by the ideal  $I_e$  defining  $J_e$  as in (17). So (cf. (10)) the hermitian sheaf defining  $h_{\Theta_e} \circ \iota_{e, P_0}$  is admissible with respect to the normalized hermitian metric

$$\mu_e := \frac{i}{2 \dim(J_e)} \sum_{f \in S_2(\Gamma_0(p))^{\text{new}}[I_e]} \frac{f \frac{dq}{q} \wedge \overline{f \frac{dq}{q}}}{\|f\|^2}. \quad (23)$$

**Remark 3.1** Notice that both  $\mu_0$  and  $\mu_e$  above are invariant by pull-back  $w_p^*$  by the Fricke involution. In particular, the Arakelov intersection products  $[\cdot, \cdot]_{\mu_0}$  and  $[\cdot, \cdot]_{\mu_e}$ , relative to  $\mu_0$  and  $\mu_e$  respectively, are  $w_p$ -invariant (as is clear from the fact that, more generally,  $w_p$  is an orthogonal symmetry on  $J_0(p)$  endowed with its quadratic form  $h_\Theta$ , which respects the orthogonal decomposition  $\prod_f J_f$  of (16)).

For any  $(1, 1)$ -form as above (even assuming milder hypothesis than smoothness if needed) one can specialize the Hodge index theorem to our modular setting (see [39], Theorem 4.16, [38], Theorem 3.26, or more generally [43], p. 85 et seq.):

**Theorem 3.2** *Let  $K$  be a number field,  $\mu$  be a (smooth)  $(1, 1)$ -form on  $X_0(p)(\mathbb{C})$ , and  $\widehat{CH}(p)_{\mathbb{R}, \mu}^{\text{num}}$  be the arithmetic Chow group with real coefficients up to numerical equivalence of  $\mathcal{X}_0(p)$  over  $\mathcal{O}_K$ , relative to  $\mu$ . Denote by  $\infty$  the horizontal divisor defined by the  $\infty$ -cusp on  $\mathcal{X}_0(p)$  over  $\mathbb{Z}$  (which is the Zariski closure of the  $\mathbb{Q}$ -point  $\infty$  in  $X_0(p)(\mathbb{Q})$ ), compactified with the normalizing condition (21). Write  $\mathbb{R} \cdot X_\infty$  for the line of divisors with real coefficients supported on some fixed full vertical fiber  $X_\infty$ . Define, for all  $v \in \text{Spec}(\mathcal{O}_K)$  above  $p$ , the  $\mathbb{R}$ -vector space:*

$$G_v := \sum_{C \neq C_\infty} \mathbb{R} \cdot C$$

where the sum runs through all the irreducible components of  $\mathcal{X}_0(p) \times_{\mathcal{O}_K} k(v)$ , except  $C_\infty$  (the one containing  $\infty(k(v))$ ). Identify finally  $J_0(p)(K)/\text{torsion}$  with the subgroups of divisors classes  $D_0$  which are compactified under the normalizing condition  $g_{D_0}(\infty) = 0$  (which is therefore different from (21)). One has a decomposition:

$$\widehat{CH}(p)_{\mathbb{R}, \mu}^{\text{num}} = (\mathbb{R} \cdot \infty \oplus \mathbb{R} \cdot X_\infty) \oplus_{v|p}^\perp G_v \oplus^\perp (J_0(p)(K) \otimes \mathbb{R}) \quad (24)$$

where the “ $\oplus^\perp$ ” mean that the direct factors are mutually orthogonal with respect to the Arakelov intersection product. Moreover, the restriction of the self-intersection product to  $J_0(p)(K) \otimes \mathbb{R}$  coincides with the opposite of the Néron-Tate pairing.

**Proof** The proof is written in [39], Theorem 4.16, for  $L_2^1$ -admissible measures (a setting allowing to define convenient actions of the Hecke algebra on the Chow group). It immediately adapts to our frame. For further computational use, we just recall how one decomposes divisors in practice. Take  $D$  in  $\widehat{CH}(p)_{\mathbb{R},\mu}^{\text{num}}$ , with degree  $d$  on the generic fiber. Then there is a vertical divisor  $\Phi_D$ , with support in fibres above places of bad reduction (that is, of characteristic  $p$ ), such that  $(D - d\infty - \Phi_D)$  has a real multiple which belongs to the neutral component  $\text{Pic}^0(\mathcal{J}_0(p))_{/\mathcal{O}_K}$ . That  $\Phi_D$  is well-defined up to multiple of full vertical fibres, so we can assume  $\Phi_D$  belongs to  $\oplus^\perp G_p$  (and is then unambiguously defined). One associates to  $(D - d\infty - \Phi_D) \in \mathbb{R} \cdot \mathcal{J}_0^0(p)(\mathcal{O}_K)$  an element  $\delta$  in  $\widehat{CH}(p)_{\mathbb{R},\mu}^{\text{num}}$  by imposing a compactification such that  $[\infty, \delta]_\mu = 0$ . The general Hodge index theorem (see for instance [43]) finally asserts that  $(D - d\infty - \Phi_D - \delta)$  can be written as an element in  $\mathbb{R} \cdot X_\infty$ .  $\square$

In order to interpret later on the Néron-Tate height (associated with some given (symmetric) invertible sheaf) as an Arakelov height in a suitable sense (see [1] paragraph 3, or [44]), we will need to compute explicitly, given  $P \in X_0(p)(K)$ , the vertical divisor  $\Phi_P = \oplus_{v|p} \Phi_{P,v}$  such that

$$[C, P - \infty - \Phi_P] = 0 \quad (25)$$

for any irreducible component of any fiber of  $\mathcal{X}_0(p) \rightarrow \text{Spec}(\mathcal{O}_K)$ , as in the proof of Theorem 3.2.

**Lemma 3.3** *Consider a bad fiber  $\mathcal{X}_0(p)_{k(v)}$ , with  $e_v$  the absolute ramification index of  $v$ , and write  $\#k(v) = p^{f_v}$ . Let  $P \in X_0(p)(K)$  and let  $C_{P,v}$  be the irreducible component of  $\mathcal{X}_0(p)_{k(v)}$  which contains  $P(k(v))$ . (Note that, as  $\mathcal{X}_0(p)$  is assumed to be regular, the section  $P$  hits each fiber on its smooth locus, so that the component  $P$  belongs to is unambiguously defined in each bad fiber.) Write  $\Phi_{P,v} = \sum_{n,m} a_{n,m} [C_{n,m}]$ , with notations as in (13). Recall that, by our convention,  $a_{C_\infty} = a_{*,0} = 0$ .*

(a) *If  $C_{P,v} = C_0$  then for all  $n, m$ ,*

$$a_{n,m} = \frac{-12}{(p-1) \cdot w_n} \cdot m.$$

(Recall  $w_n := \#\text{Aut}(s(n))/\langle \pm 1 \rangle \in \{1, 2, 3\}$ , with  $s(n)$  the supersingular point corresponding to the branch  $\{C_{n,\cdot}\}$ .)

For further use, let us henceforth write  $\Phi_{C_0}$  for the above vector  $\Phi_{P,v} \in \mathbb{Z}^{\mathcal{B}}$ .

(b) *If  $C_{P,v} = C_{n_0, m_0} \neq C_0, C_\infty$  then*

- *for  $n = n_0$  and  $m \in \{0, m_0\}$ , one has  $a_{n,m} = \left( \frac{m_0}{w_{n_0} e_v} \left( 1 - \frac{12}{(p-1)w_{n_0}} \right) - 1 \right) \cdot m$ ;*
- *for  $n = n_0$  and  $m \in \{m_0, w_{n_0} e_v\}$ , one has  $a_{n,m} = \left( \frac{m_0}{w_{n_0} e_v} \left( 1 - \frac{12}{(p-1)w_{n_0}} \right) \right) \cdot m - m_0$ ;*
- *for  $n \neq n_0$  and all  $m \in \{0, w_n e_v\}$ , one has  $a_{n,m} = \frac{-12m_0}{(p-1)w_{n_0} e_v} \cdot \frac{m}{w_n}$ .*

(c) *(Of course if  $C_{P,v} = C_\infty$  then  $\Phi_{P,v} = 0$ .)*

**Remark 3.4** We have distinguished different cases above because the proof naturally leads to doing so, and it will be of interest below to have the simpler case (a) explicitly displayed. Note however that all outputs are actually covered by the formulae of case (b). Notice also that, in case (a), all coefficients of  $\Phi_{P,v}$  satisfy

$$0 \geq a_{n,m} \geq a_0 := a_{C_0} = a_{n,w_n m} = -12e_v/(p-1).$$

As for case (b), all coefficients of  $\Phi_{P,v}$  satisfy

$$0 \geq a_{n,m} \geq a_{n_0, m_0} = \left( \frac{m_0}{w_{n_0} e_v} \left( 1 - \frac{12}{(p-1)w_{n_0}} \right) - 1 \right) \cdot m_0$$

(remember  $0 \leq m \leq w_n e_v$  for all  $m$ ). Computing the minimum of the above right-hand as a polynomial in  $m_0$  gives

$$0 \geq a_{n,m} \geq \frac{-e_v w_{n_0}}{4(1 - \frac{12}{(p-1)w_{n_0}})} \geq \frac{-e_v w_{n_0}}{4 - \frac{3}{w_{n_0}}} \geq -3e_v \quad (26)$$

(recalling we always assume  $p \geq 17$ ).

**Proof** Given the intersection matrix (15), the condition (25):  $[C, P - \infty - \Phi_{P,v}] = 0$  for all  $C$  in the fiber at  $v$  gives the matrix equation:

$$\log(\#k(v))\mathcal{M} \cdot \Phi_{P,v} = \log(\#k(v))(-1, 0, \dots, 1, 0, \dots, 0)^t \quad (27)$$

where the coefficient 1 (respectively,  $-1$ ) on the right-hand column vector is at the place corresponding to  $C_{P,v} = C_{n,m}$  (respectively, to  $C_\infty = C_{n,0}$ ) in the ordering of our component basis (13). That is however more easily solved by running through the dual graph of  $\mathcal{X}_0(p)_{k(v)}$  “branch by branch” as follows. Suppose first that  $C_{P,v} = C_0$ , and recall  $a_{C_\infty} = 0$  by convention. Equation (25) then translates into:

- for  $C = C_\infty$ :  $(-1 - \sum_{n=1}^s a_{n,1} = 0)$ ;
- for  $C = C_0$ :  $(1 + sa_0 - \sum_{n=1}^s a_{n,w_n e_v - 1} = 0)$ ;
- for all others  $C = C_{n,m}$ :  $(a_{n,m-1} - 2a_{n,m} + a_{n,m+1} = 0)$ .

The equations of the third line in turn define, for each branch (that is, for fixed  $n$ ), a sequence defined by linear double induction, with solution  $a_{n,m} = m \cdot \alpha_n$  for some  $\alpha_n$  which is easily computed to be  $\frac{-1}{w(\text{Eis}) \cdot w_n} = \frac{-12}{(p-1)w_n}$  (cf. (11)). (Note this is true even for  $e_v = 1$ ).

For case (b), the intersection equations become:

- for  $C = C_\infty$ :  $(-1 - \sum_{n=1}^s a_{n,1} = 0)$ ;
- for  $C = C_0$ :  $(sa_0 - \sum_{n=1}^s a_{n,w_n e_v - 1} = 0)$ ;
- for  $C = C_{P,v} = C_{n_0,m_0}$ :  $(1 - a_{n_0,m_0-1} + 2a_{n_0,m_0} - a_{n_0,m_0+1} = 0)$ ;
- for all others  $C = C_{n,m}$ :  $(a_{n,m-1} - 2a_{n,m} + a_{n,m+1} = 0)$ .

As above, solving these equations in all branches not containing  $C_{P,v}$  gives  $a_{n,m} = m\beta_n$  and the same is true in the branch containing  $C_{P,v}$  for  $m \in \{0, \dots, m_0\}$ . We also see that  $a_{n_0,m_0+1} = (m_0+1)\beta_{n_0} + 1$ , and then  $a_{n_0,m} = m(\beta_{n_0} + 1) - m_0$  for  $m \in \{m_0+1, w_{n_0} e_v\}$ . We have  $a_0 = w_{n_0} e_v \beta_{n_0}$  for all  $n \neq n_0$ , so let  $\beta$  be the common value of the  $\beta_n$  for  $n \neq n_0$  with  $w_n = 1$ . (There is always such an  $n$  as we assumed  $p > 13$ . Note also those computations still cover the case  $e_v = 1$ .) From  $\beta = a_0/e_v$  and  $a_0 = w_{n_0} e_v (\beta_{n_0} + 1) - m_0$  we derive

$$\beta_{n_0} = (a_0 + m_0 - w_{n_0} e_v)/w_{n_0} e_v = \frac{\beta}{w_{n_0}} + \frac{m_0}{w_{n_0} e_v} - 1.$$

Hence, because of the first equation  $(-1 - \sum_{n=1}^s a_{n,1} = 0)$ ,

$$0 = -1 - \beta_{n_0} - \sum_{1 \leq n \leq s, n \neq n_0} \beta/w_n = -\beta w(\text{Eis}) - \frac{m_0}{w_{n_0} e_v},$$

therefore

$$\beta = \frac{-m_0}{w(\text{Eis})w_{n_0} e_v} = \frac{-12 m_0}{(p-1)w_{n_0} e_v}.$$

□

Recall  $\mu$  denotes an arbitrary (smooth)  $(1, 1)$ -form on  $X_0(p)(\mathbb{C})$ .

**Lemma 3.5** (a) The class in  $\widehat{CH}(p)_{\mathbb{R}, \mu}^{\text{num}}$  of the cuspidal divisor  $(0) - (\infty)$  satisfies

$$(0) - (\infty) \equiv \Phi_{C_0}^0 := \Phi_{C_0} + \sum_{v|p} \frac{6e_v}{p-1} (\sum_C [C]) = \sum_{v|p} \sum_{n,m} \frac{6}{(p-1)} (e_v - \frac{2m}{w_n}) [C_{n,m}] \quad (28)$$

with notations as in Lemma 3.3 (a). This is an eigenvector of the Fricke  $\mathbb{Z}$ -automorphism  $w_p$  with eigenvalue  $-1$ .

(b) One has  $[\infty, \infty]_\mu = [0, 0]_\mu = [0, \infty]_\mu - \frac{6 \log p}{p-1}$ . If  $\mu$  is the Green-Arakelov measure  $\mu_0$  then  $0 \geq [\infty, \infty]_{\mu_0} = O(\log p/p)$  and similarly  $[0, \infty]_{\mu_0} = O(\log p/p)$  with  $[0, \infty]_{\mu_0}$  negative too, at least for large enough  $p$ . If  $\mu = \mu_e$  (cf. (23)) - or more generally any sub-measure of  $\mu_0$  - then  $[0, \infty]_{\mu_e} = O(p \log p)$ .

**Proof** By the Manin-Drinfeld theorem,  $(0) - (\infty)$  is torsion as a divisor in the generic fiber  $\mathcal{X}_0(p) \times_{\mathbb{Z}} \mathbb{Q}$ . One therefore has

$$(0) - (\infty) \equiv \Phi + cX_\infty$$

in the decomposition (24) of  $\widehat{CH}(p)_{\mathbb{R}, \mu}^{\text{num}}$ , for  $\Phi$  some vertical divisor with support in the fibers above  $p$ . This divisor is determined by the same equations (25) as  $\Phi_{C_0}$  in Lemma 3.3 (a). As  $\sum_{v|p} (\sum_C [C])$  is a sum of full  $v$ -fibers, it is numerically equivalent to some real multiple of the archimedean fiber  $X_\infty$ . Therefore

$$\Phi_{C_0}^0 := \Phi_{C_0} + \sum_{v|p} \frac{6e_v}{p-1} (\sum_C [C]) \equiv \Phi_{C_0} + aX_\infty$$

for some real number  $a$ . Now  $w_p$  switches the cusps  $0$  and  $\infty$  so the divisor  $(0) - (\infty)$  is anti-symmetric for  $w_p$ :

$$w_p^*((0) - (\infty)) = -((0) - (\infty))$$

and clearly  $w^*(\Phi_{C_0}^0) = -\Phi_{C_0}^0$ . The fact that  $w_p$  preserves the archimedean fiber concludes the proof of (a).

To prove (b) we compute

$$0 = [0 - \infty - \Phi_{C_0}^0, \infty]_\mu = [0, \infty]_\mu - [\infty, \infty]_\mu - \frac{6}{p-1} \log p$$

and

$$0 = [0 - \infty - \Phi_{C_0}^0, 0]_\mu = [0, 0]_\mu - [0, \infty]_\mu + \frac{6}{p-1} \log p$$

so that  $[\infty, \infty]_\mu = [0, 0]_\mu = [0, \infty]_\mu - \frac{6 \log p}{p-1}$ . The cusps  $0$  and  $\infty$  are known not to intersect on  $\mathcal{X}_0(p)_{/\mathbb{Z}}$  so that  $[0, \infty]_\mu = -g_\mu(0, \infty)$ . When  $\mu = \mu_0$ , this special value of the Arakelov-Green function has been computed by Michel and Ullmo: it satisfies

$$g_{\mu_0}(0, \infty) = \frac{1}{2g} \log p (1 + O(\frac{\log \log p}{\log p})) = O(\frac{\log p}{p})$$

by [41], formula (12) on p. 650. Finally, using [11], Theorem 7.1 (c) and paragraph 8, and plugging in Bruin's method the estimates of [41] regarding the comparison function  $F(z) = O((\log p)/p)$  between Green-Arakelov and Poincaré measures, we obtain a bound of shape  $O(p \log p)$  for  $|g_{\mu_e}(0, \infty)|$  (see also Remark 4.2). This completes the proof of (b).  $\square$

Instrumental in the sequel will be the explicit decomposition of the relative dualizing sheaf  $\omega$  in the arithmetic Chow group.

**Proposition 3.6** *The relative dualizing sheaf  $\omega$  of the minimal regular model  $\mathcal{X}_0(p) \rightarrow \mathcal{O}_K$  can be written, in the decomposition (24) of  $\widehat{CH}(p)_{\mathbb{R}, \mu_0}^{\text{num}}$  relative to the canonical Green-Arakelov  $(1, 1)$ -form  $\mu_0$ , as:*

$$\omega = (2g - 2)\infty + \sum_{v|p} \Phi_{\omega, v} + (-H_4^0 - H_3^0) + [K : \mathbb{Q}]c_{\omega}X_{\infty} \quad (29)$$

where the above components satisfy the following properties.

- The number  $c_{\omega}$  is equal to  $\frac{(1-2g)}{[K:\mathbb{Q}]}[\infty, \infty]_{\mu_0}$ , so that  $0 \leq c_{\omega} \leq O(\log p)$ .

- Set

$$H_4 := \frac{1}{2} \sum_4 (P - \frac{1}{2}(0 + \infty)), \quad H_3 := \frac{2}{3} \sum_3 (P - \frac{1}{2}(0 + \infty))$$

where the sums run over the sets (whose number of elements can be 0 or 2) of Heegner points of  $X_0(p)$  with  $j$ -invariants 1728 and 0, respectively. Then  $H_4^0 = H_4 + [K : \mathbb{Q}]c_4X_{\infty}$  and  $H_3^0 = H_3 + [K : \mathbb{Q}]c_3X_{\infty}$  for two numbers  $c_3, c_4$  with  $c_3 = O(\log p)$ , and same for  $c_4$ . (Recall this means the  $H_*$  are compacted with the normalizing condition (21), whereas the  $H_*^0$  are the orthogonal projections on  $(J_0(p)(K) \otimes \mathbb{R}) \subseteq \widehat{CH}(p)_{\mathbb{R}, \mu_0}^{\text{num}}$  of the  $H_*$ , so  $[\infty, H_*^0]_{\mu_0} = 0$ , for  $*$  = 3 or 4.)

- Finally, the component  $\Phi_{\omega, v}$  in each  $G_v$  for  $v|p$  is

$$\Phi_{\omega, v} = -12 \frac{(g-1)}{(p-1)} \sum_{n, m} \frac{m}{w_n} C_{n, m} \quad (30)$$

with notations as in (13). We therefore have  $\Phi_{\omega, v} = (g-1)\Phi_{C_0}$  using notations of Lemma 3.3. In particular, recalling  $e_v$  is the ramification index of  $K/\mathbb{Q}$  at  $v$ , the coefficients  $\omega_{n, m}$  of  $\Phi_{\omega, v}$  in (30) satisfy

$$0 \geq \omega_{n, m} \geq -e_v. \quad (31)$$

**Proof** This is available, in slightly different settings, in [41], Section 6, or [39], Section 4.4. We also crucially rely on results of Edixhoven et al. in [20].

We first estimate  $c_{\omega}$ . By Arakelov's adjunction formula,

$$-[\infty, \infty]_{\mu_0} = [\infty, \omega]_{\mu_0} = (2g - 2)[\infty, \infty]_{\mu_0} + [K : \mathbb{Q}]c_{\omega}$$

because of the orthogonality of the decomposition (24). Therefore Lemma 3.5 implies

$$0 \leq c_{\omega} = \frac{(1-2g)}{[K : \mathbb{Q}]}[\infty, \infty]_{\mu_0} = O(\log p).$$

The computations of the  $J_0(p)$ -part  $\omega_0 := -(H_3^0 + H_4^0)$  follows from the Hurwitz formula, as explained in [41], paragraph 6, p. 670. One indeed checks that, on the generic fiber  $X_0(p)_{/\mathbb{Q}} = \mathcal{X}_0(p) \times_{\mathbb{Z}} \mathbb{Q}$ , the canonical divisor is linearly equivalent to

$$(2g - 2)\infty - \left( \frac{1}{2} \sum_{j(P)=e^{i\pi/2}} ' (P - \infty) + \frac{2}{3} \sum_{j(P)=e^{2i\pi/3}} ' (P - \infty) \right)$$

where the sums  $\sum'$  are here restricted to points  $P$  at which  $X_0(p) \rightarrow X(1)$  is unramified (these are the Heegner points alluded to in our statement). It follows from the modular interpretation

that in each of those sums there are two Heegner points (if any), which are then ordinary at  $p$  (recall we assume  $p > 13 > 3$ ). Heegner points are preserved by the Atkin-Lehner involution ([25], paragraph 5, p. 90), so their specializations above  $p$  share themselves between the two components  $C_0$  and  $C_\infty$  of  $\mathcal{X}_0(p)/\mathbb{F}_p$ . This proves that the  $J_0(p)(K) \otimes_{\mathbb{Z}} \mathbb{R}$ -part of  $\omega$  is indeed  $-(H_4^0 + H_3^0)$  with  $H_4^0 = H_4 + [K : \mathbb{Q}]c_4X_\infty$  and  $H_3^0 = H_3 + [K : \mathbb{Q}]c_3X_\infty$  for some real numbers  $c_3$  and  $c_4$  (and our  $H_*$  even belong to the neutral component  $J_0(p)^0(\mathcal{O}_K)$ ). The estimates on  $c_3$  and  $c_4$  will be justified at the end of the proof.

The bad fibers divisors  $\Phi_{\omega,v} := \sum_{n,m} \omega_{n,m} [C_{n,m}]$  can be computed with the “vertical” adjunction formula ([36] Chapter 9, Theorem 1.37) as in [39], Lemma 4.22. Indeed, each irreducible component  $C$  in the  $v$ -fiber having genus 0, one has

$$[C, C + \omega] = -2 \log(\#k(v)).$$

If  $\mathcal{M}$  is the intersection matrix displayed in (15), and  $\delta_{*,*}$  is Kronecker’s symbol, we therefore have

$$C \cdot \mathcal{M} \cdot \Phi_{\omega,v} = -2 - \frac{1}{\log(\#k(v))} [C, C] - (2g - 2) \delta_{C, C_\infty} = \begin{cases} 0 & \text{if } C \neq C_\infty, C_0 \\ s - 2g & \text{if } C = C_\infty \\ s - 2 & \text{if } C = C_0 \end{cases} \quad (32)$$

that is, as  $s = g + 1$ :

$$\mathcal{M} \cdot \Phi_{\omega,v} = (g - 1)(-1, 0, \dots, 0, 1)^t.$$

That equation is (27) (up to a multiplicative scalar), which has been solved in the first case of Lemma 3.3. Therefore

$$\Phi_{\omega,v} = (g - 1) \Phi_{C_0}, \text{ that is : } \omega_{n,m} = \frac{12(1 - g)}{(p - 1)} \cdot \frac{m}{w_n}. \quad (33)$$

As noted in Remark 3.4 and using (12), this imply the coefficients  $\omega_{n,m}$  of  $\Phi_{\omega,v}$  satisfy

$$0 \geq \omega_{n,m} \geq \frac{12(1 - g)}{p - 1} e_v > -e_v.$$

We finally estimate the intersection products  $c_3 = \frac{-1}{[K:\mathbb{Q}]} [\infty, H_3]_{\mu_0}$  and  $c_4 = \frac{-1}{[K:\mathbb{Q}]} [\infty, H_4]_{\mu_0}$ . By the adjunction formula and Hriljac-Faltings’ theorem ([14], Theorem 5.1 (ii)) we compute that, for any  $P \in X_0(p)(K)$ ,

$$\begin{aligned} -2[K : \mathbb{Q}] h_\Theta(P - \frac{1}{2g-2} \omega) &= [P - \frac{1}{2g-2} \omega - \Phi_\omega(P), P - \frac{1}{2g-2} \omega - \Phi_\omega(P)]_{\mu_0} \\ &= \frac{1}{(2g-2)^2} [\omega, \omega]_{\mu_0} + \frac{g}{g-1} [P, P]_{\mu_0} - \Phi_\omega(P)^2 \end{aligned}$$

(where here  $\Phi_\omega(P)$  is a vertical divisor supported at bad fibers such that

$$[C, P - \frac{1}{2g-2} \omega - \Phi_\omega(P)] = 0 \quad (34)$$

for any irreducible component  $C$  of any bad fiber of  $\mathcal{X}_0(p)/\mathcal{O}_K$ ). Hence

$$\frac{1}{(2g-2)^2} \omega^2 + \frac{g}{g-1} [P, P]_{\mu_0} - \Phi_\omega(P)^2 = -2[K : \mathbb{Q}] h_\Theta((P - \infty) + \frac{1}{2g-2} (H_3 + H_4)). \quad (35)$$

We specialize to the case when  $P = P_*$  (where the upper star belongs to  $\{1, 2\}$  and the lower star is 4 (respectively, 3)) is one of the Heegner points occurring in  $H_4$  (respectively,  $H_3$ ). We replace for now the base field  $K$  by  $F := \mathbb{Q}(P_*) = \mathbb{Q}(\sqrt{-1})$  (respectively,  $\mathbb{Q}(\sqrt{-3})$ ). The right-hand of (35), if non-zero, is then

$$8 \log(p)(1 + o(1)) \quad (\text{respectively, } -12 \log(p)(1 + o(p))) \quad (36)$$



by [41], p. 673. If those Heegner points occur we already noticed that  $p$  splits in  $F$ , so there are two bad primes  $v, v'$  on  $\mathcal{O}_F$  (therefore two bad fibers on  $\mathcal{X}_0(p)_{/\mathcal{O}_F}$  and two  $G_v, G_{v'}$ ) to take into account. We compute  $\Phi_\omega(P_*)$  and  $\Phi_\omega(P_*)^2$ . As mentioned at the beginning of the proof,  $P_*$  specializes to the component  $C_0$  at a place, say  $v$ , of  $F$  above  $p$ , and to  $C_\infty$  at the conjugate place  $v'$ . The conditions (34) therefore give that, for any irreducible component  $C$  of the fiber at  $v$ :

$$0 = [C, P_*^* - \frac{1}{2g-2}\omega - \Phi_\omega(P_*)_v] = [C, 0 - \infty - \frac{1}{2g-2}\Phi_{\omega,v} - \Phi_\omega(P_*)_v],$$

and using Lemma 3.3, Lemma 3.5 and (33) one obtains

$$\Phi_\omega(P_*)_v = -\frac{1}{2g-2}\Phi_{\omega,v} + \Phi_{C_0,v} = \frac{1}{2}\Phi_{C_0,v}$$

whereas, at  $v'$ :

$$\Phi_\omega(P_*)_{v'} = -\frac{1}{2g-2}\Phi_{\omega,v'} = -\frac{1}{2}\Phi_{C_0,v'}.$$

Using Lemmas 3.3 and 3.5 again we therefore have

$$\Phi_\omega(P_*)^2 = \sum_{w|p} \frac{1}{4}\Phi_{C_0,w}^2 = \sum_{w|p} \frac{1}{4}[\Phi_{C_0,w}, 0 - \infty] = \frac{1}{2}a_0 \log p = -\frac{6 \log(p)}{p-1}. \quad (37)$$

As for the self-intersection of  $\omega$ , one knows that

$$\omega_{\mathcal{X}_0(p)/\mathbb{Z}}^2 = 3g \log(p)(1 + o(1))$$

(cf. [60], Introduction), so for the dualizing sheaf  $\omega_{\mathcal{X}_0(p)/\mathcal{O}_F}$  of  $\mathcal{X}_0(p)$  over  $\mathcal{O}_F$  (instead of  $\mathbb{Z}$ ) one has  $\omega^2 = 6g \log(p)(1 + o(1))$  (see Remark 3.7 below). Summing-up, equation (35) implies that

$$[P_*, P_*^*]_{\mu_0} = O(\log(p)) \quad (38)$$

for each Heegner point  $P_*$ . Now, on the other hand, the vertical divisor  $\Phi_{P_*^*}$  in the sense of Lemma 3.3 is  $\Phi_{P_*^*} = \Phi_{C_0,v}$  for  $v$  the place of  $F$  where  $P_*$  specialize on  $C_0$  and not  $C_\infty$ . Therefore

$$\begin{aligned} -4h_\Theta(P_*^* - \infty) &= [P_*^* - \infty - \Phi_{P_*^*}, P_*^* - \infty - \Phi_{P_*^*}]_{\mu_0} \\ &= -2[P_*^*, \infty]_{\mu_0} + [P_*^*, P_*^*]_{\mu_0} + [\infty, \infty]_{\mu_0} + (\Phi_{P_*^*})^2 - 2[P_*^*, \Phi_{P_*^*}] \\ &= -2[P_*^*, \infty]_{\mu_0} + [P_*^*, P_*^*]_{\mu_0} + [\infty, \infty]_{\mu_0} - (\Phi_{P_*^*})^2 \end{aligned} \quad (39)$$

whence, using (36), (37), (38) and Lemma 3.5(b):

$$[P_*^*, \infty]_{\mu_0} = \frac{1}{2}([P_*^*, P_*^*]_{\mu_0} + [\infty, \infty]_{\mu_0} - (\Phi_{C_0,v})^2 + 4h_\Theta(P_*^* - \infty)) = O(\log p).$$

Putting everything together and using Lemma 3.5 once more, we conclude that

$$c_4 = -\frac{1}{[K:\mathbb{Q}]}[\infty, H_*]_{\mu_0} = \frac{1}{2[K:\mathbb{Q}]}(-[\infty, P_4^1 + P_4^2]_{\mu_0} + [\infty, 0 + \infty]_{\mu_0}) = O(\log p) \quad (40)$$

similarly for  $c_3$ . (Note that the Arakelov intersection products, in the computations around (39), were over  $F = \mathbb{Q}(P_*)$  and not  $K$ , although we did not indicate this in the notations in order to try keeping them not too heavy. We however want quantities over  $K$  for the statement of the theorem, so we need considering Arakelov products over  $K$  in (40) above.)  $\square$

**Remark 3.7** Let us justify (the well-known fact) that, as used in the above proof, the quantity  $\frac{1}{[F:K]}[\omega]^2$  is independent of the number field extension  $F/K$ . Let  $X \rightarrow \text{Spec}(\mathcal{O}_K)$  be a *semistable* curve over the ring of integers of any number field  $K$ , and  $F/K$  be any extension. We

know that the relative dualizing sheaf  $\omega_{X/\mathcal{O}_K}$  commutes with base-change ([59], p. 277; or more precisely [36], Theorem 6.4.32 together with Theorem 6.4.9 (b) and Proposition 10.3.15 (b)). However, the base change  $X \times_{\mathcal{O}_K} \mathcal{O}_F$  is not quite the model  $X_{/\mathcal{O}_F}$  we want to work with, because it is not regular in general, as required by classical Arakelov theory: we need make a few blowing-ups on  $X \times_{\mathcal{O}_K} \mathcal{O}_F$  to obtain our regular minimal model  $X_{/\mathcal{O}_F}$ , as recalled above. There is however a method of Mumford and Bost, explained in [9], paragraph 2.4 and 5.3, to define a lift  $\tilde{\mathcal{F}}$  of any invertible hermitian sheaf  $\mathcal{F}$  via the desingularization map  $X_{/\mathcal{O}_F} \rightarrow X \times_{\mathcal{O}_K} \mathcal{O}_F$ , which in turns allows to extend Arakelov pairings to any arithmetic surface which is only assumed to be *normal*. So if  $\mathcal{F}$  and  $\mathcal{G}$  are elements of the arithmetic Chow group  $\widehat{CH}(p)_{\mathbb{R},\mu}^{\text{num}}$  of  $X_{/\mathcal{O}_K}$  up to numerical equivalence, then

$$[\tilde{\mathcal{F}}, \tilde{\mathcal{G}}]_{\mathcal{O}_F} = [F : K][\mathcal{F}, \mathcal{G}]_{\mathcal{O}_K}.$$

Doing so we obtain two hermitian sheaves  $\tilde{\omega}_K \times \mathcal{O}_F$  and  $\omega_F$  on  $X_F$ , which might differ at their vertical components above bad fibers. Now one readily checks that equations of type (32) above (allowing to define the vertical components of our true  $\omega_F$ ) are the same as those given by Bost conditions (that is true for any curve, not only our modular ones). In other words,  $\tilde{\omega}_K \times \mathcal{O}_F$  and  $\omega_F$  on  $X_{/\mathcal{O}_F}$  define the same elements in  $\widehat{CH}(p)_{\mathbb{R},\mu}^{\text{num}}$ . In particular  $\frac{1}{[F:K]}[\omega]^2$  is independent of the field extension  $F/K$ .

In our explicit modular case, it is an easy computation to check directly the latter fact, given the explicit description of Proposition 3.6. (That is not a circular reasoning, as in doing so we use only property (30) of the Proposition.)

## 4 $j$ -height and $\Theta$ -height

We have recalled in Section 2.2 the definition of the naive  $j$ -height on  $X_0(p)$ . We aim to compare this with the height  $h_\Theta$  induced by some well-chosen embedding of  $X_0(p)$  in its jacobian. As a first step towards this, we start by making a comparison of  $h_j$  with the “degree component” (in the sense of Theorem 3.2) of the hermitian sheaf  $\omega$ .

**Proposition 4.1** *Let  $h_j$  be Weil’s  $j$ -height on  $X_0(p)$ , and let  $\mu_0$  and  $\mu_e$  be the  $(1, 1)$ -forms defined in (22) and (23). Recall  $\sup_{X_0(p)(\mathbb{C})} g_\mu$  stands for the upper bound for all Green functions  $g_{\mu,a}$  relative to some point  $a$  of  $X_0(p)(\mathbb{C})$  (and to the measure  $\mu$ ).*

*There are real numbers  $C_0$  and  $C_e$  such that if  $p$  is a prime number,  $K$  is a number field and  $P$  belongs to  $X_0(p)(K)$  then*

$$\begin{aligned} h_j(P) &\leq \frac{(p+1)}{[K:\mathbb{Q}]} \left( [P, \infty]_{\mu_0} + [K:\mathbb{Q}] \sup_{X_0(p)(\mathbb{C})} g_{\mu_0} + O(1) \right) \\ &\leq \frac{(p+1)}{[K:\mathbb{Q}]} [P, \infty]_{\mu_0} + C_0 \cdot p^3 \end{aligned} \tag{41}$$

and similarly

$$\begin{aligned} h_j(P) &\leq \frac{(p+1)}{[K:\mathbb{Q}]} \left( [P, \infty]_{\mu_e} + [K:\mathbb{Q}] \sup_{X_0(p)(\mathbb{C})} g_{\mu_e} + O(1) \right) \\ &\leq \frac{(p+1)}{[K:\mathbb{Q}]} [P, \infty]_{\mu_e} + C_e \cdot p^3. \end{aligned} \tag{42}$$

**Remark 4.2** The functions  $C_0 \cdot p^3$  and  $C_e \cdot p^3$  come from the main result of [11], which states explicitly that the suprema of our functions verify:

$$\sup_{X_0(p)(\mathbb{C})} g_{\mu_0} \leq 0.088 \cdot p^2 + 7.7 \cdot p + 1.6 \cdot 10^4, \tag{43}$$

cf. [11], Theorem 1.2 for the case of  $\sup_{X_0(p)(\mathbb{C})} g_{\mu_0}$ ). It follows from measures comparison (cf. (49) below) and the method of P. Bruin in loc. cit. that this holds for  $\sup_{X_0(p)(\mathbb{C})} g_{\mu_e}$  too. Actually, it seems that, at least in the case of  $X_0(p)$ , if we plug in Bruin's method the estimates of [41] regarding the comparison function  $F(z)$  between Green-Arakelov and Poincaré measures, we even obtain a bound of shape  $O(p \log p)$  instead of  $O(p^2)$  (cf. [11], p. 263, and Paragraph 8 with references therein (Theorem 7.1 in particular)). The same again holds true for the Green function  $g_{\mu_e}$ . As suggested by P. Autissier, estimates of the same order  $O(p \log p)$  could also be proven from results in [61]. The main theorems of [31] and [3] might even yield that the functions  $O(p^2)$  of (43) could be replaced by a uniform bound  $O(1)$ .

**Proof** This is essentially a question of measure comparisons on  $X_0(p)(\mathbb{C})$ , between  $j^*(\mu_{FS})$  on one hand (where  $\mu_{FS}$  is the Fubini-Study (1,1)-form on  $X(1)(\mathbb{C}) \simeq \mathbb{P}^1(\mathbb{C})$ ) and the Green-Arakelov form  $\mu_0$  (respectively,  $\mu_e$ ) on the other hand. We adapt the main result of [19].

Define first a somewhat canonical Arakelov intersection product  $[\cdot, \cdot]_{\mu_{FS}}$  on the projective line using  $\mu_{FS}$ . We write  $\mathbb{P}_{\mathcal{O}_K}^1 = \text{Proj}(\mathcal{O}_K[x_0, x_1]) = \overline{\text{Spec}}^{\text{Zar}}(\mathcal{O}_K[j])$  (with  $j = x_1/x_0$ ), so that the horizontal divisor  $\infty(\mathcal{O}_K)$  is  $V(x_0)$  and, writing  $P = [x_0 : x_1]$ , its associate Green function is:

$$g_{\mu_{FS}, \infty}(P) = g_{\mu_{FS}, \infty}(j(P)) = \frac{1}{2} \log \left( \frac{|x_0|^2}{|x_0|^2 + |x_1|^2} \right) = -\frac{1}{2} \log(1 + |j(P)|^2)$$

at any point different from  $\infty = [0 : 1]$ . Then for any  $P$  in  $X(1)(K)$  one has

$$\left| h_j(P) - \frac{1}{[K : \mathbb{Q}]} [j(P), \infty]_{\mu_{FS}} \right| \leq \frac{1}{2} \log(2).$$

Applying [19], Theorem 9.1.3 and its proof to the setting described above gives, for any  $P$  in  $X_0(p)(K)$ ,

$$[j(P), \infty]_{\mu_{FS}} \leq [P, j^*(\infty)]_{\mu_0} + (p+1) \sum_{\sigma} \sup_{X_0(p)_{\sigma}} g_{\mu_0} + \frac{1}{2} \sum_{\sigma} \int_{X_0(p)_{\sigma}} \log(|j|^2 + 1) \mu_0 \quad (44)$$

where  $\sigma$  runs through the infinite places of  $K$  and  $X_0(p)_{\sigma} := X_0(p) \times_{\mathcal{O}_K, \sigma} \mathbb{C}$ .

We estimate the right-hand terms of (44). As mentioned above, Theorem 1.2 of [11] provides a universal  $c$  such that

$$\sup_{X_0(p)(\mathbb{C})} g_{\mu_0} \leq c p^2. \quad (45)$$

To estimate the last integrals of (44) we recall that, on the union of disks of ray  $|q| < r$  around the cusps (that is, on the image in  $X_0(p)(\mathbb{C})$  of the open subset  $D_r := \{z \in \mathcal{H}, \Im(z) > -(\log r)/2\pi\}$  in Poincaré upper-half plane  $\mathcal{H}$ ), for some fixed  $r$  in  $]0, 1[$ , one has

$$\left| \frac{f(q)}{q} \right| \leq \frac{2}{(1-r)^2}$$

for any newform  $f$  in  $S_2(\Gamma_0(p))$ . (See for instance [20], Lemma 11.3.7 and its proof). We also know that the Peterson norm of such an  $f$  verifies  $\|f\|^2 \geq \pi e^{-4\pi}$  (see [20], Lemma 11.1.2). Choose  $r = 1/2$  to fix ideas. On  $D_{1/2}$ , we therefore have

$$\mu_0 \leq \frac{64e^{4\pi}}{\pi} \frac{i}{2} dq \wedge \overline{dq}.$$

(Sharper bounds should be affordable, but the one above is good enough for our present purpose.) It follows that there exists some real  $A$  such that, in the decomposition

$$\int_{X_0(p)(\mathbb{C})} \log(|j|^2 + 1) \mu_0 = \int_{X_0(p)(\mathbb{C}) \cap D_{1/2}} \log(|j|^2 + 1) \mu_0 + \int_{X_0(p)(\mathbb{C}) \setminus D_{1/2}} \log(|j|^2 + 1) \mu_0 \quad (46)$$

the first term of the right-hand side satisfies

$$\int_{X_0(p)(\mathbb{C}) \cap D_{1/2}} \log(|j|^2 + 1) \mu_0 \leq \frac{64e^{4\pi}}{\pi} [\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(p)] \int_{X(1)(\mathbb{C}) \cap D_{1/2}} \log(|j|^2 + 1) \frac{i}{2} dq \wedge \overline{dq} \leq (p+1)A.$$

As for the second term, remembering that  $\mu_0$  has total mass 1 on  $X_0(p)(\mathbb{C})$  we check that

$$\int_{X_0(p)(\mathbb{C}) \setminus D_{1/2}} \log(|j|^2 + 1) \mu_0 \leq M_{1/2} := \max_{X(1)(\mathbb{C}) \setminus D_{1/2}} (\log(|j|^2 + 1))$$

whence the existence of some absolute real number  $A_0$  such that

$$\int_{X_0(p)(\mathbb{C})} \log(|j|^2 + 1) \mu_0 \leq (p+1)A_0. \quad (47)$$

There is therefore a constant  $C$  for which

$$h_j(P) \leq \frac{1}{[K : \mathbb{Q}]} [P, j^*(\infty)]_{\mu_0} + (p+1) \left( \sup_{X_0(p)(\mathbb{C})} g_{\mu_0} + A_0 \right) \leq \frac{1}{[K : \mathbb{Q}]} [P, j^*(\infty)]_{\mu_0} + C \cdot p^3.$$

With notations of Lemma 3.5, one further has

$$j^*(\infty) = p(0) + (\infty) \equiv (p+1)\infty + p \cdot \Phi_{C_0}^0 \quad (48)$$

as elements of  $\widehat{CH}(p)_{\mathbb{R}, \mu_0}^{\mathrm{num}}$ . Using Lemma 3.5 (a) we obtain

$$|[P, \Phi_{C_0}^0]| \leq [K : \mathbb{Q}] \frac{6 \log p}{p-1}$$

so that

$$h_j(P) \leq \frac{1}{[K : \mathbb{Q}]} [P, (p+1)\infty]_{\mu_0} + C_0 \cdot p^3$$

which is (41).

The proof of (42) proceeds along the same lines, with one more ingredient. Applying Theorem 9.1.3 of [19] with the measure  $\mu_e$  instead of  $\mu_0$  gives the corresponding version of (44). To obtain an upper bound for  $\sup_{X_0(p)(\mathbb{C})} g_{\mu_e}$  we recall that the theorem of Kowalski-Michel (20) asserts that  $\dim(J_e) \geq \dim(J_0(p))/5$  for large enough  $p$ . Our measure  $\mu_e := \frac{1}{\dim(J_e)} \sum_{S_e} \frac{i}{2} \frac{f \frac{dq}{q} \wedge \overline{f \frac{dq}{q}}}{\|f\|^2}$  (cf. (23)) therefore satisfies

$$0 \leq \mu_e \leq \frac{g}{\dim(J_e)} \mu_0 \leq 5\mu_0. \quad (49)$$

This shows that Bruin's theorem ([11], Theorem 7.1) again provides a universal  $c_e$  such that

$$\sup_{X_0(p)(\mathbb{C})} g_{\mu_e} \leq c_e p^2. \quad (50)$$

Using (47) we obtain:

$$\int_{X_0(p)(\mathbb{C})} \log(|j|^2 + 1) \mu_e \leq (p+1)A_e. \quad (51)$$

Finally, equivalence (48) remains naturally true in the Chow group  $\widehat{CH}(p)_{\mathbb{R}, \mu_e}^{\mathrm{num}}$  relative to the measure  $\mu_e$  (and actually, any measure) instead of  $\mu_0$ , as remarked in Lemma 3.5 (a). This completes the proof of (42).  $\square$

We now relate  $h_j$  and the Néron-Tate height  $h_\Theta$  relative to the  $\Theta$ -divisor (cf. (8)) and a convenient embedding of the curve inside its jacobian:

**Theorem 4.3** *There are real numbers  $\gamma, \gamma_1$  such that the following holds. Let  $K$  a number field and  $p$  a prime number. Let  $\omega^0 := -(H_4 + H_3)$  be the 0-component of the canonical sheaf  $\omega$  on  $X_0(p)$  over  $K$  (cf. Proposition 3.6). If  $P$  is a point of  $X_0(p)(K)$  then*

$$h_\Theta(P - \infty + \frac{1}{2}\omega^0) = \frac{1}{[K : \mathbb{Q}]}[P, g \cdot \infty]_{\mu_0} + O(\log p) \quad (52)$$

so that

$$h_j(P) \leq (12 + o(1)) \cdot h_\Theta(P - \infty + \frac{1}{2}\omega^0) + \gamma \cdot p^3 \quad (53)$$

(where here  $(12 + o(1)) = 12(p+1)/(p-13)$ ) and

$$h_j(P) \leq (48 + o(1)) \cdot h_\Theta(P - \infty) + \gamma_1 \cdot p^3. \quad (54)$$

**Remark 4.4** Theorem 4.3 offers only one sense of inequality between  $j$ -height and  $\theta$ -height: with our method of proof, it is harder to give an effective form to the reverse inequality, because of the metrics comparisons we use (see below). One could however probably achieve this by using variants of the results of Merkl used here, as will be alluded in the next section, cf. Proposition 6.4 and around.

**Proof** Recall  $\mathcal{X}_0(p)$  denotes the minimal regular model of  $X_0(p)$  on  $\text{Spec}(\mathcal{O}_K)$ , that  $\mathcal{J}_0(p)$  is the Néron model of  $J_0(p)$  on the same basis, and  $\mathcal{J}_0(p)^0$  denotes its neutral component. Let  $\delta$  be an element of  $J_0(p)(K)$ . We shall denote by  $\tilde{\delta} = \delta + \Phi_\delta$  (for  $\Phi_\delta$  some vertical divisor (with coefficient 0 on the component containing  $\infty$ , following our conventions)) the associate element of the neutral component  $\mathcal{J}_0(p)^0(\mathcal{O}_K)$  (that is, the one that has degree zero on each irreducible component, in any fiber, of  $\mathcal{X}_0(p)$ , and therefore defines a point of  $\mathcal{J}_0(p)^0(\mathcal{O}_K)$ ). For any point  $P$  in  $X_0(p)(K) \hookrightarrow \mathcal{X}_0(p)(\mathcal{O}_K)$ , let also  $\Phi_P$  be a vertical divisor on  $\mathcal{X}_0(p)$ , with support on the bad fibers, such that  $(P - \infty - \Phi_P)$  has divisor class belonging to the neutral component  $\text{Pic}^0(\mathcal{J}_0(p))_{/\mathcal{O}_K}$ . (Note that our notation  $\Phi_P$  is consistent with (25).) As usual, imposing  $\Phi_P$  has everywhere trivial  $\infty$ -component determines it uniquely and we can compute it explicitly as in Lemma 3.3. Write  $\Phi_P = \sum_{v \in M_K, v|p} \sum_{C_v} a_{C_v} [C_v]$ , where the sum is taken on irreducible components  $C_v$  of vertical bad fibers of  $\mathcal{X}_0(p)$ . Using notations of Lemma 3.3 (b) we also define the following new vertical divisor at bad fibers:

$$\Phi_\Theta := \sum_{v \in M_K, v|p} \sum_{Q_v} a_{C_{Q_v}} C_{Q_v} = \sum_{v|p} \sum_{(n_0, m_0)} a_{n_0, m_0}^v C_{n_0, m_0} \quad (55)$$

(so that

$$a_{n_0, m_0}^v = \left( \frac{m_0}{w_{n_0} e_v} \left( 1 - \frac{12}{(p-1)w_{n_0}} \right) - 1 \right) \cdot m_0.$$

Our very definitions imply

$$\Phi_P^2 = [P, \Phi_P] = [P, \Phi_\Theta]. \quad (56)$$

By Faltings' Hodge index theorem, the Néron-Tate height  $h_\Theta(P - \infty + \delta)$  is:

$$\begin{aligned} h_\Theta(P - \infty + \delta) &= \frac{-1}{2[K : \mathbb{Q}]} [P - \infty + \tilde{\delta} - \Phi_P, P - \infty + \tilde{\delta} - \Phi_P]_{\mu_0} \\ &= \frac{1}{2[K : \mathbb{Q}]} ([P, \omega + 2\infty - 2\tilde{\delta}]_{\mu_0} + 2[P, \Phi_P]_{\mu_0} - [\Phi_P, \Phi_P]_{\mu_0} \\ &\quad + [\tilde{\delta}, 2\infty + 2\Phi_P - \tilde{\delta}]_{\mu_0} - [\infty, \infty]_{\mu_0}) \\ &= \frac{1}{2[K : \mathbb{Q}]} ([P, \omega + 2\infty - 2\tilde{\delta} + \Phi_\Theta]_{\mu_0} + [\tilde{\delta}, 2\infty - \tilde{\delta}]_{\mu_0} - [\infty, \infty]_{\mu_0}) \\ &= \frac{1}{[K : \mathbb{Q}]} [P, \tilde{\omega}_\delta]_{\mu_0} \end{aligned} \quad (57)$$

with  $\tilde{\omega}_\delta := \left(\frac{1}{2}(\omega + \Phi_\Theta) + \infty - \tilde{\delta}\right) + c_\delta X_\infty$  for  $X_\infty$  some fixed archimedean fiber of  $\mathcal{X}_0(p)$  and  $c_\delta$  is the real number:

$$c_\delta = \frac{1}{2} \left( -[\infty, \infty]_{\mu_0} + [\tilde{\delta}, 2\infty - \tilde{\delta}]_{\mu_0} \right). \quad (58)$$

Note that  $\tilde{\omega}_\delta$  does not depend on  $P$  (certainly  $\Phi_\Theta$  was introduced to that aim).

Let us now take  $\delta = \omega^0/2 = -(H_3 + H_4)/2 \in \frac{1}{2} \cdot J_0(p)^0(\mathbb{Q})$ , as defined in Proposition 3.6. (This is Riemann's characteristic (the “ $\kappa$ ” of [28], p. 138 for instance, that is the generic fiber of the  $J_0(p)(\mathbb{Q}) \otimes \mathbb{R}$ -part of  $\omega$  in the decomposition (29).) Then

$$\tilde{\omega}_\Theta := \tilde{\omega}_\delta = \left( g \cdot \infty + \frac{1}{2}(\Phi_\omega + \Phi_\Theta) + c_\Theta X_\infty \right) \quad (59)$$

for  $c_\Theta$  which, still using notations of Proposition 3.6 and its proof, is explicitly given by:

$$\begin{aligned} \frac{1}{[K : \mathbb{Q}]} c_\Theta &= \frac{1}{2} \left( c_\omega - c_4 - c_3 + \frac{1}{2} h_\Theta(H_3 + H_4) - \frac{1}{[K : \mathbb{Q}]} ([\infty]_{\mu_0}^2 + [\infty, H_3 + H_4]_{\mu_0}) \right) \\ &= \frac{1}{2} \left( c_\omega - \frac{1}{[K : \mathbb{Q}]} [\infty]_{\mu_0}^2 + \frac{1}{2} h_\Theta(H_3 + H_4) \right). \end{aligned}$$

As in the proof of Proposition 3.6 we invoke p. 673 of [41] to assert  $h_\Theta(H_3 + H_4) = O(\log(p))$ . We moreover know from the same Proposition and from Lemma 3.5 that both  $|c_\omega| = O(\log p)$  and  $[\infty, \infty]_{\mu_0} = [K : \mathbb{Q}] O(\log p/p)$ , so that

$$c_\Theta = [K : \mathbb{Q}] O(\log p). \quad (60)$$

The contribution of  $(\Phi_\omega + \Phi_\Theta)$  is controlled by Lemma 3.3 and Remark 3.4: on one hand,

$$\begin{aligned} 0 \geq [P, \Phi_\Theta] &= [P, \Phi_P] = \sum_{v \in M_K, v|p} a_{C_P, v} \log(\#k_v) \geq \sum_{v \in M_K, v|p} -3e_v \log(p^{f_v}) \\ &\geq -3[K : \mathbb{Q}] \log(p) \end{aligned} \quad (61)$$

(where as usual  $e_v$  and  $f_v$  are the ramification index and the residual degree respectively of  $K$  at  $v$ ). On the other hand, by (31), the coefficients of the vertical components  $\Phi_{\omega, v}$  satisfy  $0 \geq \omega_{n, m} \geq -e_v$ , so denoting by  $\omega_{n_P, m_P, v}$  the coefficient in  $\Phi_{\omega, v}$  of the component containing  $P(k(v))$  we have:

$$0 \geq [P, \Phi_\omega] = \sum_{v|p} \omega_{n_P, m_P, v} \log(\#k(v)) \geq \sum_{v|p} -e_v \log(p^{f_v}) = -[K : \mathbb{Q}] \log(p). \quad (62)$$

Putting (60), (61) and (62) together we obtain:

$$\begin{aligned} h_\Theta(P - \infty + \frac{\omega^0}{2}) &= \frac{1}{[K : \mathbb{Q}]} [P, g \cdot \infty + \frac{1}{2}(\Phi_\omega + \Phi_\Theta) + c_\Theta X_\infty]_{\mu_0} \\ &= \frac{1}{[K : \mathbb{Q}]} [P, g \cdot \infty]_{\mu_0} + O(\log p) \end{aligned} \quad (63)$$

and eventually, using Proposition 4.1 (and (12)),

$$h_j(P) \leq 12 \frac{p+1}{p-13} h_\Theta(P - \infty + \frac{1}{2} \omega^0) + O(p^3).$$

The last estimate (54) of the theorem just comes the fact that  $h_\Theta$  is a quadratic form and that

$$h_\Theta(\omega^0) = O(\log p) \quad (64)$$

by the results of [41] now many times mentioned.  $\square$

**Remark 4.5** The above proof builds an explicit element

$$\tilde{\omega}_\Theta = \left( g \cdot \infty + \frac{1}{2}(\Phi_\omega + \Phi_\Theta) + c_\Theta X_\infty \right) \quad (65)$$

of  $\widehat{CH}(p)_{\mathbb{R}, \mu_0}^{\text{num}}$  such that, for any  $P \in X_0(p)(K)$  one has  $h_\Theta(P - \infty + \frac{1}{2}\omega^0) = \frac{1}{[K:\mathbb{Q}]}[P, \tilde{\omega}_\Theta]_{\mu_0}$  and the error terms given by  $\Phi_\omega$ ,  $\Phi_\Theta$  and  $c_\Theta$  are controlled using (60), (61) and (62) as in (63). This immediately generalizes to any (regular) curve embedded in its jacobian, and therefore gives another proof that the height induced by the Theta height (on the jacobian) *over fields of some fixed degree* is an arakelovian height on *all* the curve (that is, for all  $\overline{\mathbb{Q}}$ -points which specialize to the smooth locus of the curve on fields of the above mentioned degree), and not only the points specializing to the component mapped to the jacobian's neutral component.

On the other hand, coming back to the case of  $X_0(p)$ , the involution  $w_p$  acts as an isometry (actually, an orthogonal symmetry) with respect to the quadratic form  $h_\Theta$  on  $J_0(p)(K) \otimes_{\mathbb{Z}} \mathbb{R}$ . Indeed,  $w_p$  acts as multiplication by  $\pm 1$  on each factor of Shimura's decomposition up to isogeny:

$$J_0(p) \sim \prod_{f \in G_{\mathbb{Q}} \cdot S_2(\Gamma_0(p))^{\text{new}}} J_f$$

whose factors are  $h_\Theta$ -orthogonal subspaces. (See also [38], Corollaire 4.3, or [39], Theorem 4.5 (3).) As  $w_p(\omega^0) = \omega^0$  (cf. the proof of Proposition 3.6), this implies

$$h_\Theta(P - \infty + \frac{1}{2}\omega^0) = h_\Theta(w_p(P - \infty + \frac{1}{2}\omega^0)) = h_\Theta(w_p(P) - 0 + \frac{1}{2}\omega^0) = h_\Theta(w_p(P) - \infty + \frac{1}{2}\omega^0)$$

(using once more that  $(0) - (\infty)$  is torsion), so that

$$[P, \tilde{\omega}_\Theta]_{\mu_0} = [w_p(P), \tilde{\omega}_\Theta]_{\mu_0} = [P, w_p^*(\tilde{\omega}_\Theta)]_{w_p^*(\mu_0)} = [P, w_p^*(\tilde{\omega}_\Theta)]_{\mu_0} \quad (66)$$

(cf. Remark 3.1). This suggests to write  $\tilde{\omega}_\Theta$  in a  $w_p$ -eigenbasis of  $\widehat{CH}(p)_{\mathbb{R}, \mu}^{\text{num}}$  instead of that of Theorem 3.2, for instance

$$\widehat{CH}(p)_{\mathbb{R}, \mu_0}^{\text{num}} = \mathbb{R} \cdot \frac{1}{2}(0 + \infty) \oplus \mathbb{R} \cdot X_\infty \oplus_{v|p} \Gamma_v \oplus (J_0(p)(K) \otimes \mathbb{R}) \quad (67)$$

where now the  $\Gamma_v$  decomposes as the direct sum of eigenspaces  $\Gamma_v^{w_p=-1}$  and  $\Gamma_v^{w_p=+1}$ , with basis:

$$\{C_{n,m}^- := C_{n,m} - w_p(C_{n,m})\}_{\substack{1 \leq n \leq s \\ 0 \leq m \leq e w_n / 2}} \text{ and } \{C_{n,m}^+ := C_{n,m} + w_p(C_{n,m}) - C_0 - C_\infty\}_{\substack{1 \leq n \leq s \\ 1 \leq m \leq e w_n / 2}} \quad (68)$$

respectively. Using Lemma 3.5 and Proposition 3.6, a somewhat lengthy but easy computation allows to check that now

$$\tilde{\omega}_\Theta = g \cdot \frac{1}{2}(0 + \infty) + \Phi_\Theta^+ + \gamma_\Theta X_\infty$$

where  $\Phi_\Theta^+$  is an explicit vertical divisor above  $p$  with  $w_p^*(\Phi_\Theta^+) = \Phi_\Theta^+$ , so that indeed

$$w_p^*(\tilde{\omega}_\Theta) = \tilde{\omega}_\Theta$$

(thus recovering (66)).

Consider for instance the case of  $\mathcal{X}_0(p)$  over  $\mathbb{Z}$ , for  $p \equiv 1 \pmod{12}$  (that is,  $\mathcal{X}_0(p)_{/\mathbb{Z}}$  is regular, so that there is no need to blow-up singular points of width larger than 1). Here  $\Gamma_v = \Gamma_v^- = \mathbb{R} \cdot C_0^- = \mathbb{R} \cdot ([C_\infty] - [C_0])$  and one readily checks that

$$\tilde{\omega}_\Theta = \frac{g}{2}(0 + \infty) + \gamma_\Theta X_\infty \quad (69)$$

that is, there is no  $\Gamma_v$ -component at all in that case. Evaluating  $h_\Theta(\frac{1}{2}\omega^0)$  as in the proof of Proposition 3.6 and using Lemma 3.5,

$$\gamma_\Theta = -\frac{g}{2}[\infty, 0 + \infty]_{\mu_0} + h_\Theta(\frac{1}{2}\omega^0) = gO(\log p/p) + O(\log p) = O(\log p).$$

## 5 Height of modular curves, and the various $W_d$ 's

For later applications of the explicit arithmetic Bézout theorem displayed in next section (Proposition 6.1), we now estimate the degree and Faltings height of the image of  $X_0(p)$ , together with its various  $d^{\text{th}}$ -symmetric products (commonly called “ $W_d$ ”), within either the jacobian  $J_0(p)$  or its quotient  $J_e$ , relative to the  $\Theta$ -polarization. We actually estimate those heights both in the normalized Néron-Tate sense, and for some good (“Moret-Bailly”) projective models, to be defined shortly.

Instrumental to us will here be Zhang’s control of heights in terms of essential minima (cf. [62], Theorem 5.2), that is, the two inequalities:

$$\mu_{\mathcal{F}}^{\text{ess}}(\mathcal{X}) \leq \frac{h_{\mathcal{F}}(\mathcal{X})}{\deg_{\mathcal{F}}(\mathcal{X})} \leq (\dim \mathcal{X}) \mu_{\mathcal{F}}^{\text{ess}}(\mathcal{X}) \quad (70)$$

for  $(\mathcal{X}, \mathcal{F})$  a pair of an arithmetic scheme - that is, a scheme  $\mathcal{X}$  which is projective and flat over, say, some ring of algebraic integers  $\mathcal{O}_K$ , and  $\mathcal{X}$  is endowed with an admissible hermitian invertible sheaf  $\mathcal{F}$ .

If  $(\mathcal{X}, \mathcal{F})$  is such a model over  $\mathcal{O}_K$  of a polarized abelian variety  $(X, F)$  over  $K = \text{Frac}(\mathcal{O}_K)$ , and  $Y$  is a cycle in the generic fiber  $X$ , recall that its normalized Néron-Tate height relative to  $F$  is defined as the limit

$$\hat{h}_F(Y) := \lim_{n \rightarrow \infty} \frac{1}{N^{2n}} h_{\mathcal{F}}(\overline{[N]_* Y})$$

where  $N$  is any fixed integer larger than 1 and  $\overline{[N]_* Y}$  denotes the schematic closure in  $\mathcal{X}$  of the image  $[N]_* Y$  of  $Y$  by multiplication by  $N$  in the generic fiber  $X = \mathcal{X}_K$ . This normalized height, which is a direct generalization of the classical notion of Néron-Tate height for points, is known not to depend neither on the model  $\mathcal{X}$  of  $X$ , nor the extension  $\mathcal{F}$  of  $F$ , nor its hermitian structure (and not on  $N$ ), so that the notation  $\hat{h}_F(\cdot)$  is unambiguous. We refer to [1], Proposition-Définition 3.2 of Section 3 and references therein for more details. We will actually use the extension of the two inequalities (70) to the case where the heights and essential minima are those given by the limit process defining Néron-Tate heights, cf. Théorème 3.4 of [1].

### 5.1 Néron-Tate heights

**Proposition 5.1** *Let  $X$  be the image via  $\pi_A \circ \iota_{\infty}: X_0(p) \rightarrow A$  of the modular curve  $X_0(p)$  mapped to a quotient  $\pi_A: J_0(p) \rightarrow A$  of its jacobian, endowed with the polarization  $\Theta_A$  induced by the  $\Theta$ -divisor (see (2), (7) and around). The degree and normalized Néron-Tate height of  $X$  satisfy:*

$$\deg_{\Theta_A}(X) = \dim(A) = O(p)$$

and

$$\hat{h}_{\Theta_A}(X) = O(p \log p).$$

(Of course this definition implicitly assumes the existence of some projective model for  $J_0(p)$ , whose mere existence is clear. “Good” explicit examples will be built below.)

**Proof** If  $(A, \Theta_A) = (\text{Jac}(X_0(p)), \Theta)$ , it is well-known that the  $\Theta$ -degree of  $X_0(p)$  (or any curve, actually), embedded in its jacobian via some any Albanese embedding, equals its genus. That can be seen in many ways, among which one can invoke Wirtinger’s theorem ([24], p. 171), which actually yields the desired result for any quotient  $(A, \Theta_A)$ :

$$\deg_{\Theta_A}(X) = \int_X \omega_A = \int_{X_0(p)} \iota_{\infty}^* \pi_A^*(\omega) = \int_{X_0(p)} \sum_{f \in S_2^{\text{new}}(\Gamma_0(p))[I_A]} \frac{i f \frac{dq}{q} \wedge \overline{f \frac{dq}{q}}}{2 \|f\|^2} = \dim A \leq g(X_0(p)).$$

We then apply again the fact (12) that the genus  $g(X_0(p))$  is roughly  $p/12$ . (We could also have more simply invoked the fact that the degree is decreasing by projection, as in the proofs below.)



Similarly, the height of  $X$  is, by definition, the arithmetic intersection product  $(c_1(\Theta_A)^2|X|_{\mathcal{O}_K})$ . It is however simpler to use Zhang's formula (70). Indeed, it follows from the main result of [41] that the (first) essential minimum  $\mu_{\Theta}^{\text{ess}}(X_0(p))$  is  $O(\log p)$ . As the height decreases by projection (cf. (4), paragraph 2.1.2), the same is true for  $\mu_{\Theta_A}^{\text{ess}}(X) = O(\log p)$ , and (70) gives

$$\hat{h}_{\Theta_A}(X) = O(p \log p). \quad \square$$

Now for the Néron-Tate normalized height of symmetric squares and variants:

**Proposition 5.2** *Assume  $X := X_0(p)$  has gonality strictly larger than 2 (which is true as soon as  $p > 71$ , cf. [48]). Let  $\iota := \iota_{\infty}: X_0(p) \hookrightarrow J_0(p)$  be the Albanese embedding as in Proposition 5.1. Let  $X^{(2)}$  be the symmetric square  $X_0(p)^{(2)}$  embedded in  $J_0(p)$  via  $(P_1, P_2) \mapsto \iota(P_1) + \iota(P_2)$ , and similarly let  $X^{(2),-}$  be the image of  $(P_1, P_2) \mapsto \iota(P_1) - \iota(P_2)$ . Let  $X_{e^{\perp}}^{(2)}$  (respectively,  $X_{e^{\perp}}^{(2),-}$ ) be the projection of  $X^{(2)}$  (resp.,  $X^{(2),-}$ ) to  $J_e^{\perp}$  (the “orthogonal complement” to the winding quotient  $J_e$ , see paragraph 2.2.3). Then, with notations as in Proposition 5.1 (taking  $A = J_0(p)$  and  $A = J_e^{\perp}$  respectively), one has*

$$\deg_{\Theta}(X^{(2)}) = O(p^2) = \deg_{\Theta}(X^{(2),-}), \quad \hat{h}_{\Theta}(X^{(2)}) = O(p^2 \log p) = \hat{h}_{\Theta}(X^{(2),-})$$

and the same holds with the quotient objects:

$$\deg_{\Theta_e^{\perp}}(X_{e^{\perp}}^{(2)}) = O(p^2) = \deg_{\Theta_e^{\perp}}(X_{e^{\perp}}^{(2),-}); \quad \hat{h}_{\Theta_e^{\perp}}(X_{e^{\perp}}^{(2)}) = O(p^2 \log p) = \hat{h}_{\Theta_e^{\perp}}(X_{e^{\perp}}^{(2),-}).$$

**Proof** Denoting by  $p_1$  and  $p_2$  the obvious projections we factor in the common way (cf. [46], paragraph 3, Proposition 1 on p. 320) our maps (over  $\mathbb{Q}$ ) as follows:

$$\begin{array}{ccccc} & & & & A \\ & & & & \nearrow p_2 \\ X_0(p) \times X_0(p) & \xrightarrow{\pi_A \iota \times \pi_A \iota} & A \times A & \xrightarrow{M} & A \times A \\ & & (x, y) & \mapsto & (x + y, x - y) \\ & & & & \searrow p_1 \\ & & & & A \end{array} \quad (71)$$

so  $X^{(2)} = p_1 \circ M \circ (\iota \times \iota)(X_0(p) \times X_0(p))$  and  $X^{(2),-} = p_2 \circ M \circ (\iota \times \iota)(X_0(p) \times X_0(p))$  when  $A = J_0(p)$ , and same with  $X_{e^{\perp}}^{(2)}$  and  $X_{e^{\perp}}^{(2),-}$  after composing with  $\pi_A$  for  $A := J_e^{\perp}$ . We endow  $A \times A$  with the hermitian sheaf  $\Theta_A^{\boxtimes 2}$ . Then  $M^*(\Theta_A^{\boxtimes 2}) \simeq (\Theta_A^{\boxtimes 2})^{\otimes 2}$  ([46], p. 320). Therefore, writing  $X$  for  $\pi_A \iota(X_0(p))$  in short and using Proposition 5.1,

$$\deg_{\Theta_A^{\boxtimes 2}}(M(X \times X)) = 4 \deg_{\Theta_A^{\boxtimes 2}}(X \times X) = 8(\deg_{\Theta_A}(X))^2 = O(g^2).$$

As degree decreases by our projections and  $O(g^2) = O(p^2)$ ,  $\deg_{\Theta_A}(X^{(2)})$  and  $\deg_{\Theta_A}(X^{(2),-})$  are  $O(p^2)$ .

Using (70) (for the Néron-Tate normalized heights) and obvious notations,

$$\mu_{\Theta_A^{\boxtimes 2}}^{\text{ess}}(X \times X) \leq 2\mu_{\Theta_A}^{\text{ess}}(X).$$

This implies that  $\mu_{\Theta_A^{\boxtimes 2}}^{\text{ess}}(M(X \times X)) \leq 4\mu_{\Theta_A}^{\text{ess}}(X)$ . Invoking (70) again, inequalities (20) and Proposition 5.1 together with the fact that the height of points also decreases by projection,

$$\mu_{\Theta_A}^{\text{ess}}(X^{(2)}) \leq \mu_{\Theta_A^{\boxtimes 2}}^{\text{ess}}(M(X \times X)) \leq 4\mu_{\Theta_A}^{\text{ess}}(X) \leq 4 \frac{\hat{h}_{\Theta_A}(X)}{\deg_{\Theta_A}(X)} \leq O(\log p).$$

Therefore

$$\hat{h}_{\Theta_A}(X^{(2)}) = O(p^2 \log p)$$

and the same holds true for the other variants of Proposition 5.2 (at least if  $\deg_{\Theta_A}(X)$  is larger than something linear in  $p$ , which is indeed the case (cf. (20)) when  $A = J_e^{\perp}$ ).  $\square$

(Note that this proof applies more generally to any sub-quotient of  $J_0(p)$ , taking into account the issue of dimensions ratios in the final formulas.)

## 5.2 Moret-Bailly models and associate projective heights

To build the projective models of the jacobian (over  $\mathbb{Z}$ , or finite extensions) that we shall need for our arithmetic Bézout, we use Moret-Bailly theory as follows. For much more about similar constructions in the general setting of abelian varieties, we refer to [8], 2.4 and 4.3; see also [50].

Let therefore  $(J, L(\Theta))$  stand for the principally polarized abelian variety  $J_0(p)$  endowed with the invertible sheaf associate with its symmetric theta divisor, defined over some small extension of  $\mathbb{Q}$  (cf. (79) below and around for more details). Endow the complex base-changes of the associate invertible sheaf  $L(\Theta)$  with its cubist hermitian metric. If  $\mathcal{N}_{J, \mathcal{O}_K}$  is the Néron model of  $J$  over the ring of integers  $\mathcal{O}_K$  of a number field  $K$ , we know this is a semistable scheme over  $\mathcal{O}_K$ , whose only non-proper fibers are above primes  $\mathfrak{P}$  of characteristic  $p$  (where it then is purely toric). At any such  $\mathfrak{P}$ , with ramification index  $e_{\mathfrak{P}}$ , the group scheme  $\mathcal{N}_{J, \mathcal{O}_K}$  has components group

$$\Phi_{\mathfrak{P}} \simeq (\mathbb{Z}/N_0 e_{\mathfrak{P}} \mathbb{Z}) \times (\mathbb{Z}/e_{\mathfrak{P}} \mathbb{Z})^{g-1} \quad (72)$$

for  $g := \dim J$  and  $N_0 := \text{num}(\frac{p-1}{12})$  (cf. e.g. [35], Proposition 1.1.a)). We choose and fix an integer  $N > 0$  and a number field  $K \supseteq \mathbb{Q}(J[2N])$ , for all this paragraph, such that all the  $2N$ -torsion points in  $J$  have values in  $K$ . One then observes from (72) that  $2N$  divides all the ramification indices  $e_{\mathfrak{P}}$ , and Proposition II.1.2.2 on p. 45 of [44] asserts that  $L(\Theta)$  has a cubist extension, let us denote it by  $\mathcal{L}(\Theta)$ , to the open subgroup scheme  $\mathcal{N}_{J, N}$  of  $\mathcal{N}_{J, \mathcal{O}_K}$  whose fibers have components group killed by  $N$ . Such an extension  $\mathcal{L}(\Theta)$  is actually symmetric ([44], Remarque II.1.2.6.2) and uniquely determined (Théorème II.1.1.i) on p. 40 of loc. cit.). Moreover  $\mathcal{L}(\Theta)$  is ample on  $\mathcal{N}_{J, N}$  ([44], Proposition VI.2.1 on p. 134). Its powers  $\mathcal{L}(\Theta)^{\otimes r}$  are even very ample on  $\mathcal{N}_{J, N} \times_{\mathcal{O}_K} \mathcal{O}_K[1/2p]$  as soon as  $r \geq 3$  (as follows from the general theory of theta functions). Provided  $N > 1$ , the sheaf  $\mathcal{L}(\Theta)^{\otimes N}$  is spanned by its global sections on the whole of  $\mathcal{N}_{J, N}$  ([44], Proposition VI.2.2).

Picking-up a basis of *generic* global sections in  $H^0(J_0(p)_K, L(\Theta)^{\otimes N})$  we thus defines a map  $J_0(p)_K \xrightarrow{j_N} \mathbb{P}_K^n$ , for  $n = N^g - 1$ . Assume our generic global sections extend to a set  $\mathcal{S}$  of  $H^0(\mathcal{N}_{J, N}, \mathcal{L}(\Theta)^{\otimes N})$ . Let  $\mathcal{J}$  be the schematic closure in  $\mathbb{P}_{\mathcal{O}_K}^n$  of the generic fiber via the associate composed embedding  $J_K \hookrightarrow \mathbb{P}_K^n \hookrightarrow \mathbb{P}_{\mathcal{O}_K}^n$ . Let  $\mathcal{M} := (\sum_{s \in \mathcal{S}} \mathcal{O}_K \cdot s)$  be the subsheaf of  $\mathcal{L}(\Theta)$  spanned by  $\mathcal{S}$  on  $\mathcal{J}$ . Let  $\nu: \tilde{\mathcal{N}}_{J, N} \rightarrow \mathcal{N}_{J, N}$  be the blowing-up at base points for  $\mathcal{M}$  on  $\mathcal{N}_{J, N}$ . We have a commutative diagram

$$\begin{array}{ccccc} & & \tilde{\mathcal{N}}_{J, N} & & \\ & \nearrow & \downarrow \iota_{\mathcal{N}} & \searrow^{j_N} & \\ J_K & \hookrightarrow & \mathcal{J} & \xrightarrow{j} & \mathbb{P}_{\mathcal{O}_K}^n. \end{array} \quad (73)$$

Considering the complex base-changes of the generic fiber we note that  $\mathcal{M}$  is automatically endowed with a cubist hermitian structure induced by that of  $L(\Theta)_{\mathbb{C}}$ .

**Definition 5.3** *Given an integer  $N > 1$ , the “good model” for  $(J_0(p), L(\Theta)^{\otimes N})$  relative to the set  $\mathcal{S}$  in  $H^0(\mathcal{N}_{J, N}, \mathcal{L}(\Theta)^{\otimes N})$  will be the projective scheme  $\mathcal{J}$  over  $\text{Spec}(\mathcal{O}_K)$  enhanced with the hermitian sheaf  $\mathcal{M} := (\sum_{s \in \mathcal{S}} \mathcal{O}_K \cdot s)$  constructed above.*

Outside of the base points for  $\mathcal{M}$  on  $\mathcal{N}_{J, N}$  we have

$$\iota_{\mathcal{N}}^* \mathcal{M} = j_N^* \mathcal{O}_{\mathbb{P}_{\mathcal{O}_K}^n}^*(1) \simeq \mathcal{L}(\Theta)^{\otimes N} \quad (74)$$

so we insist those “good models” for  $(J_0(p), L(\Theta)^{\otimes N})$  will indeed compute ( $N$  times) the Néron-Tate height of *certain*  $\mathbb{Q}$ -points (those whose closure factorizes through  $\mathcal{N}_{J, N}$  deprived from the base points for  $\mathcal{S}$ ), but definitely *not all* algebraic points. For general points, still, one can deduce from the work of Bost ([8], 4.3) the following inequality.

**Proposition 5.4** *For any point  $P$  in  $J_0(p)(\overline{\mathbb{Q}})$ , let  $\varepsilon_P$  be the section of  $\mathcal{J}(\overline{\mathbb{Z}})$  it defines by Zariski closure in the model of Definition 5.3. Let  $h_{\mathcal{M}}(P) = \frac{1}{[\mathbb{Q}(P):\mathbb{Q}]} \widehat{\deg}(\varepsilon_P^*(\mathcal{M}))$  be the normalized projective height of  $P$  relative to  $\mathcal{M}$ . Then one has*

$$h_{\mathcal{M}}(P) \leq N \hat{h}_{\Theta}(P).$$

**Proof** We briefly sum-up [8], 2.4 and 4.3, using our above notations. (Note that this statement has nothing to see with modular jacobians, and holds for any abelian variety over a number field.) Let  $N'$  be some integer such that  $\varepsilon_P$  belongs to  $\mathcal{N}_{J,N'}(\mathcal{O}_F)$  for some ring of integers  $\mathcal{O}_F$ . Up to replacing  $\mathcal{O}_F$  by a sufficiently ramified finite extension, we can assume  $L(\Theta)$  has a cubist extension  $\mathcal{L}(\Theta)$  to all of  $\mathcal{N}_{J,N'}$  over  $\mathcal{O}_F$  (cf. [44], Proposition II.1.2.2). One can therefore compute

$$\hat{h}_\Theta(P) = \frac{1}{N} \frac{1}{[F:\mathbb{Q}]} \widehat{\deg}(\varepsilon_P^*(\mathcal{L}(\Theta)^{\otimes N})).$$

As in (73) however we see that there is no well-defined map from  $\mathcal{N}_{J,N'}$  to  $\mathbb{P}^n$ , because  $\mathcal{L}(\Theta)^{\otimes N}$  needs not be spanned by elements of  $\mathcal{S}$  on all of  $\mathcal{N}_{J,N'}$  (even though it of course is on the generic fiber). The way to remedy this is to replace  $\mathcal{N}_{J,N'}$  by its blowing-up  $\nu: \tilde{\mathcal{N}}_{J,N'} \rightarrow \mathcal{N}_{J,N'}$  at base points for  $\mathcal{M}$ . Now we do have a map  $f: \tilde{\mathcal{N}}_{J,N'} \rightarrow \mathcal{J} (\hookrightarrow \mathbb{P}^n)$  such that the Zariski closure of  $f_*(\tilde{\mathcal{N}}_{J,N'})$  identifies with the base-change of  $\mathcal{J}$  to  $\mathcal{O}_F$ . We moreover have on  $\tilde{\mathcal{N}}_{J,N'}$  that

$$f^*(\mathcal{M}) = \nu^*(\mathcal{M}) \otimes \mathcal{O}(-E)$$

where  $E$  is the exceptional divisor of the blow-up (which is effective by definition). The section  $\varepsilon_P$  of  $\mathcal{N}_{J,N'}$  lifts to some  $\tilde{\varepsilon}_P$  of  $\tilde{\mathcal{N}}_{J,N'}$ , and

$$\begin{aligned} h_{\mathcal{M}}(P) &= \frac{1}{[F:\mathbb{Q}]} \widehat{\deg}(\varepsilon_P^*(\mathcal{M})) = \frac{1}{[F:\mathbb{Q}]} \widehat{\deg}(\tilde{\varepsilon}_P^*(f^*(\mathcal{M}))) \\ &\leq \frac{1}{[F:\mathbb{Q}]} \widehat{\deg}(\tilde{\varepsilon}_P^*(\nu^*(\mathcal{M}))) = \frac{1}{[F:\mathbb{Q}]} \widehat{\deg}(\varepsilon_P^*(\mathcal{L}(\Theta)^{\otimes N})) = N \hat{h}_\Theta(P). \end{aligned}$$

□

The following straightforward generalization to higher dimension will be useful in next section.

**Corollary 5.5** *If  $V$  is an irreducible variety in  $J_0(p)$  over a number field  $K$ , of dimension  $d_V$ , and  $\mathcal{V}$  its Zariski closure in the model of Definition 5.3, then*

$$h_{\mathcal{M}}(\mathcal{V}) \leq (d_V + 1) N^{d_V+1} \hat{h}_\Theta(V).$$

**Proof** Combine Zhang's formula (70) with Proposition 5.4. □

Recall from (6) that one can define the “pseudo-projection”  $\mathcal{P}_{\tilde{J}_{e^\perp}}(\iota_\infty(X_0(p)))$  of the image of  $X_0(p) \xrightarrow{\iota_\infty} J_0(p)$  on the subabelian variety  $\tilde{J}_{e^\perp} \subseteq J_0(p)$ . Let  $X_{e^\perp}$  be any of its irreducible components, and  $\mathcal{X}_{e^\perp}$  its Zariski closure in  $\mathcal{J}$  over  $\overline{\mathbb{Z}}$ . Define more generally by  $\mathcal{X}^{(2)}$ ,  $\mathcal{X}^{(2),-}$ ,  $\mathcal{X}_{e^\perp}^{(2)}$  and  $\mathcal{X}_{e^\perp}^{(2),-}$  as the Zariski closures in  $\mathcal{J}$  of the relevant objects  $X^{(2)}, \dots$ , as were defined in Proposition 5.2.

Note that, by construction, the degree and normalized Néron-Tate height of  $X_{e^\perp}$  (and other similar pseudo-projections:  $X_{e^\perp}^{(2)}$  etc.), as an irreducible subvariety of  $J_0(p)$  endowed with  $\hat{h}_\Theta$ , are those of  $\pi_{J_e^\perp}(X_0(p)) = X_{e^\perp}^{(2),-}$  relative to the only natural hermitian sheaf of  $J_e^\perp$ , that is, the  $\Theta_e^\perp = \Theta_{J_e^\perp}$  described in paragraph 2.1.2 and estimated in Proposition 5.1.

**Corollary 5.6** *For any fixed positive multiple  $N$  of  $N_0 = \text{num}(\frac{p-1}{12})$ , let  $(\mathcal{J}, \mathcal{M})$  be the model for  $(J_0(p), L(\Theta)^{\otimes N})$  given in Definition 5.3. Let  $\mathcal{X}_0(p)$  be the Zariski closure in  $\mathcal{J}$  of the image of  $X_0(p) \xrightarrow{\iota_\infty} J_0(p)$ , and denote more generally by  $\mathcal{X}^{(2)}$ ,  $\mathcal{X}^{(2),-}$ ,  $\mathcal{X}_{e^\perp}^{(2)}$  and  $\mathcal{X}_{e^\perp}^{(2),-}$  the Zariski closures in this model of the relevant objects  $X^{(2)}, \dots$ , defined in Proposition 5.2. Then their  $\mathcal{M}^{\otimes \frac{1}{N}}$ -heights are bounded by similar functions as their Néron-Tate avatars (Proposition 5.2). Explicitly,  $h_{\mathcal{M}^{\otimes \frac{1}{N}}}(\mathcal{X}_0(p))$  is at most  $O(p \log p)$ , and the heights  $h_{\mathcal{M}^{\otimes \frac{1}{N}}}(X^{(2)}), \dots, h_{\mathcal{M}^{\otimes \frac{1}{N}}}(X_{e^\perp}^{(2),-})$  are all bounded from above by some  $O(p^2 \log p)$ .*

**Proof** Combine Zhang's formula (70) with Propositions 5.1, 5.2 and 5.4.  $\square$

### 5.3 Explicit modular version of Mumford's repulsion principle

We conclude this section by writing-down, for later use, an explicit version of Mumford's well-know "repulsion principle" for points, in the case of modular curves.

**Proposition 5.7** *For  $P$  and  $Q$  two different points of  $X_0(p)(\overline{\mathbb{Q}})$  one has*

$$\hat{h}_\Theta(P - Q) \geq \frac{g-2}{4g} \left( \hat{h}_\Theta(P - \infty) + \hat{h}_\Theta(Q - \infty) \right) - O(p^2). \quad (75)$$

**Proof** Let  $K$  be a number field such that both  $P$  and  $Q$  have values in  $K$ . Using notations of Section 3, the adjunction formula and Hodge index theorem give

$$\begin{aligned} 2[K : \mathbb{Q}] \hat{h}_\Theta(P - Q) &= -[P - Q - \Phi_P + \Phi_Q, P - Q - \Phi_P + \Phi_Q]_{\mu_0} \\ &= [P + Q, \omega]_{\mu_0} + 2[P, Q]_{\mu_0} + [\Phi_P - \Phi_Q]^2 \\ &\geq [P + Q, \omega]_{\mu_0} - 2[K : \mathbb{Q}] \sup g_{\mu_0} + [\Phi_P - \Phi_Q]^2. \end{aligned}$$

In the same way,

$$\begin{aligned} [P, \omega]_{\mu_0} &= 2[K : \mathbb{Q}] \hat{h}_\Theta(P - \infty) - 2[P, \infty]_{\mu_0} + [\infty]_{\mu_0}^2 - [\Phi_P]^2 \\ &\geq [K : \mathbb{Q}] \hat{h}_\Theta(P - \infty + \frac{1}{2}\omega^0) - 2[P, \infty]_{\mu_0} + [\infty]_{\mu_0}^2 - [\Phi_P]^2 \end{aligned}$$

where the last inequality comes from the quadratic nature of  $\hat{h}_\Theta$ , plus the fact that the error term of (75) allows us to assume  $\hat{h}_\Theta(P - \infty) \geq \frac{1}{(12-8\sqrt{2})} \hat{h}_\Theta(\omega^0) = O(\log p)$  (cf. (64), end of proof of Theorem 4.3). Now, by (52),

$$\hat{h}_\Theta(P - \infty + \frac{1}{2}\omega^0) = \frac{1}{[K : \mathbb{Q}]} [P, g \cdot \infty]_{\mu_0} + O(\log p)$$

and using Lemma 3.5 and Remark 3.4 gives

$$[P, \omega]_{\mu_0} \geq \frac{g-2}{g} [K : \mathbb{Q}] \hat{h}_\Theta(P - \infty + \frac{1}{2}\omega^0) + [K : \mathbb{Q}] O(\log p).$$

As  $[\Phi_P, \Phi_Q] = [P, \Phi_Q] = [Q, \Phi_P]$ , we have  $||[\Phi_P, \Phi_Q]|| \leq 3[K : \mathbb{Q}] \log p$  by Remark 3.4 again. Using Remark 4.2 about  $\sup g_{\mu_0}$ , and putting everything together, we obtain

$$\hat{h}_\Theta(P - Q) \geq \frac{g-2}{2g} \left( \hat{h}_\Theta(P - \infty + \frac{1}{2}\omega^0) + \hat{h}_\Theta(Q - \infty + \frac{1}{2}\omega^0) \right) - O(p^2)$$

which, by our previous remarks, can again be written as

$$\hat{h}_\Theta(P - Q) \geq \frac{g-2}{4g} \left( \hat{h}_\Theta(P - \infty) + \hat{h}_\Theta(Q - \infty) \right) - O(p^2). \quad \square$$

(For large  $p$ , the angle between two points of large enough and equal height is here therefore at least  $2 \arcsin(1/2\sqrt{2}) \geq \pi/5$ . Of course the natural value is  $\pi/2$ , to which one tends when sharpening the computations.)

## 6 Arithmetic Bézout theorem with cubist metric

We display in this section an explicit version of Bézout arithmetic theorem, in the sense of Philippon or Bost-Gillet-Soulé ([52], [10]), for intersections of cycles in our modular abelian varieties over number fields, with the following variants: we use Arakelovian heights (in the sense of Faltings as in [21], see also [1]) on higher-dimensional cycles, and we endow the implicit hermitian sheaf for this height with its cubist metric (instead of Fubini-Study).

It indeed seems one generally uses Fubini-Study metrics for arithmetic Bézout because they are the only natural explicit ones available on a general projective space (which of course is an (almost) necessary frame for Bézout-like statements). They moreover have the pleasant feature that the relevant projective embeddings have tautological basis of global sections with sup-norm less than 1, which allows for instance proving at hand that the induced Faltings height is non-negative on effective cycles (cf. [21], Proposition 2.6). For our present purposes however, we need to obtain bounds eventually involving the Néron-Tate heights of points, that is, Arakelov heights induced by cubist metrics. One could in principle have tried working with Fubini-Study metrics as in [10] and then directly compare with Néron-Tate heights; but comparison terms tend to be huge: in the case of rational points, for instance (that is, horizontal cycles of relative dimension 0), within jacobians, those error terms are bounded by Manin and Zahrin ([37]) linearly in the ambient projective dimension, which is exponential in the dimension of the abelian variety (that is, for our modular curves, exponential in the level). It is therefore much preferable to stick to cubist metrics. This implies we avoid the use of joins as in [10], as those need a natural sheaves metrization on the whole of the ambient projective spaces: we instead use plain Segre embeddings. The numerical extra-charge to pay essentially consists in binomial coefficients popping-up on the picture, which do not significantly alter the quantitative bounds we obtain.

We also need working with projective models which are “almost” compactifications of relevant Néron models of our modular jacobians. This will be done with bits of Moret-Bailly theory introduced in the above Section 5.

Let us also recall that there still is another approach for such arithmetic Bézout theorems which uses Chow forms ([52], [54]). That is however known to amount working with Faltings height relative to Fubini-Study metrics again ([52]-I, [58]).

Finally, regarding generality: it would of course be desirable to have a proof available for arbitrary abelian varieties. But many of the present arguments are quite particular to our application to  $J_0(p)$ . We therefore preferred here to work in this precise setting from the beginning, instead of considering a somewhat artificial generality.

The main result for this section is:

**Proposition 6.1** *Let  $(J_0(p), \Theta)$  be defined over the number field  $K$ , endowed with the principal and symmetric polarization  $\Theta$ . Let  $V, W$  be two  $K$ -subvarieties of  $J_0(p)$ , irreducible of dimension  $d_V := \dim_K V$  and  $d_W := \dim_K W$  respectively, such that*

$$d_V + d_W \leq g = \dim J_0(p),$$

*and assume  $V \cap W$  has dimension 0.*

*If  $P$  is an element of  $(V \cap W)(K)$ , then its Néron-Tate  $\Theta$ -height satisfies*

$$\begin{aligned} \hat{h}_\Theta(P) \leq & \frac{1}{2}(4N_0^2)^{d_V+d_W} \left[ (d_W + 1) \binom{d_V + d_W + 1}{d_V} \hat{h}_\Theta(W) \deg_\Theta(V) \right. \\ & \left. + (d_V + 1) \binom{d_V + d_W + 1}{d_W} \hat{h}_\Theta(V) \deg_\Theta(W) \right] \\ & + O(p^3)(d_V + d_W) \frac{1}{2}(4N_0^2)^{d_V+d_W-1} \binom{d_V + d_W}{d_V} \deg_\Theta(V) \deg_\Theta(W) \end{aligned} \quad (76)$$

*and the  $O(p^3)$  error term above can be replaced by  $O(p \log p)$  under Autissier’s result in [4] (cf. Lemma 6.6 below).*

**Remark 6.2** Our arguments will show that (76) actually holds when  $\hat{h}_\Theta(P)$  is replaced by  $\sum_i \hat{h}_\Theta(P_i)$ , with  $\cup_i P_i$  the whole support of  $V \cap W$ .

Let us describe the strategy of proof, which occupies the rest of this section. We henceforth fix a prime number  $p$ , define the integer  $N := 4N_0^2 = 4 \text{num}^2(\frac{p-1}{12})$ , and recall the Moret-Bailly projective model  $(\mathcal{J}, \mathcal{M})$  of  $(J_0(p), L(\Theta)^{\otimes N})$  given by Definition 5.3, relative to this  $N$  and to some given set of global sections  $\mathcal{S}$  in  $H^0(\mathcal{N}_{J,N}, \mathcal{L}(\Theta)^{\otimes N})$ , to be described later. That model is defined over some  $\mathcal{O}_K$ . Consider the morphisms:

$$\begin{array}{ccc} \mathcal{J} & \xrightarrow{\Delta} & \mathcal{J} \times \mathcal{J} \\ & & \mathcal{P} \downarrow \\ & & \mathbb{P}_{\mathcal{O}_K}^n \times \mathbb{P}_{\mathcal{O}_K}^n \end{array} \quad \begin{array}{c} \searrow \iota \\ \xrightarrow{S} \end{array} \quad \mathbb{P}_{\mathcal{O}_K}^{n^2+2n} \quad (77)$$

where  $\Delta$  is the diagonal map,  $n = N^g - 1$ ,  $\mathcal{P}$  is the product of two  $\mathcal{S}$ -embeddings  $\mathcal{J} \xrightarrow{\mathcal{J}} \mathbb{P}^n = \mathbb{P}_{\mathcal{O}_K}^n$  and the application  $\iota: \mathcal{J} \times \mathcal{J} \rightarrow \mathbb{P}^{n^2+2n}$  is the composition of the Segre embedding  $S$  with  $\mathcal{P}$ . As sheaves,

$$S^*(\mathcal{O}_{\mathbb{P}^{n^2+2n}}(1)) = \mathcal{O}_{\mathbb{P}^n}(1) \otimes_{\mathcal{O}_K} \mathcal{O}_{\mathbb{P}^n}(1)$$

and

$$\mathcal{P}^*(\mathcal{O}_{\mathbb{P}^n}(1) \otimes_{\mathcal{O}_K} \mathcal{O}_{\mathbb{P}^n}(1)) = \mathcal{M} \otimes_{\mathcal{O}_K} \mathcal{M} =: \mathcal{M}^{\boxtimes 2}$$

so that

$$\iota^*(\mathcal{O}_{\mathbb{P}^{n^2+2n}}(1)) = \mathcal{M}^{\boxtimes 2}$$

and

$$\Delta^* \iota^* \mathcal{O}_{\mathbb{P}^{n^2+2n}}(1) = \mathcal{M} \otimes_{\mathcal{O}_J} \mathcal{M} = \mathcal{M}^{\boxtimes 2}. \quad (78)$$

We naturally endow the sheaves  $\mathcal{M}^{\boxtimes 2}$ ,  $\mathcal{M}^{\otimes 2}$ , and so on, with the hermitian structures induced by the cubist metric on the various  $\mathcal{M}_\sigma$  for  $\sigma: K \hookrightarrow \mathbb{C}$ , denoted by  $\|\cdot\|_{\text{cub}}$ .

We then pick two copies  $(x_i)_{0 \leq i \leq n}$  and  $(y_j)_{0 \leq j \leq n}$  of the canonical basis of global sections for each  $\mathcal{O}_{\mathbb{P}^n}(1)$  on the two factors of  $\mathbb{P}_{\mathcal{O}_K}^n \times \mathbb{P}_{\mathcal{O}_K}^n$  of (77), which give our basis  $\mathcal{S}$  by restriction to  $\mathcal{J}$ . Then we provide the sheaf  $\mathcal{O}_{\mathbb{P}^{n^2+2n}}(1)$  on  $\mathbb{P}_{\mathcal{O}_K}^{n^2+2n}$  with the basis of global sections  $(z_{i,j})_{0 \leq i,j \leq n}$ , each of which is mapped to  $x_i \otimes_{\mathcal{O}_K} y_j$  under  $S^*$ . Define  $\mathcal{D}$  as the diagonal linear subspace of  $\mathbb{P}_{\mathcal{O}_K}^{n^2+2n}$  defined by the *linear* equations  $z_{i,j} = z_{j,i}$  for all  $i, j$ .

Let  $V, W \subseteq J$  be two closed subvarieties over  $K$ . The support of  $V \cap W$  is the same as that of  $(\iota \circ \Delta)^{-1}(\mathcal{D} \cap \iota(V \times W))$ . To bound from above the height of points in  $V \cap W$  it is therefore sufficient to estimate Faltings' height of  $\mathcal{D} \cap \iota(V \times W)$ , relative to the hermitian line bundle  $\mathcal{O}_{\mathbb{P}^{n^2+2n}}(1)|_{\iota(V \times W)}$  endowed with the cubist metric. As  $\mathcal{D}$  is a linear subspace, that height is the same as that of  $(V \times W)$ , up to an explicit error term depending on the degree. This error term is a priori linear in the number of (relevant) equations for  $\mathcal{D}$  - and this is way too high. But if one knows  $V \cap W$  has dimension 0, it is enough to choose  $(\dim V + \dim W)$  equations (up to perhaps increasing a bit the size of the set whose height we estimate), which makes the error term much better.

That is the basic strategy of proof for Proposition 6.1. To make it effective however we must control the “error terms” alluded to in the preceding lines, and those depend crucially on the supremum of the set  $\mathcal{S}$  of global sections describing the projective embedding  $\mathcal{J} \hookrightarrow \mathbb{P}_{\mathcal{O}_K}^n$ . We build that  $\mathcal{S}$  using theta functions, as follows.

As  $J_0(p)$  is principally polarized (over  $\mathbb{Q}$ ), the complex extension of scalars  $J_0(p)(\mathbb{C})$  can be given a classical complex uniformization  $\mathbb{C}^g / (\mathbb{Z}^g + \tau \mathbb{Z}^g)$  for some  $\tau$  in Siegel's upper half plane. The associate Riemann theta function:

$$\theta(z) = \sum_{m \in \mathbb{Z}^g} \exp(i\pi^t m \cdot \tau \cdot m + 2i\pi^t m \cdot z) \quad (79)$$

defines the tautological global section 1 of a trivialization of  $\mathcal{O}_{J_0(p)}(\Theta_{\mathbb{C}}) (= \mathcal{M}_{\mathbb{C}}^{\otimes 1/N})$  for  $\Theta_{\mathbb{C}}$  some  $W_{g-1}$ . More precisely, Riemann's classical results (e.g. [24], Theorem on p. 338) assert that  $\text{div}(\theta(z)) = \Theta_{\mathbb{C}}$  is the divisor with support  $\{\kappa_{P_0} + \sum_{i=1}^{g-1} \iota_{P_0}(P_i), P_i \in X_0(p)(\mathbb{C})\}$ , where for any  $P_0 \in X_0(p)(\mathbb{C})$  we write  $\iota_{P_0}: X_0(p) \hookrightarrow J_0(p)$  for the Albanese morphism with base point  $P_0$ , and  $\kappa = \kappa_{P_0} = \frac{\iota_{P_0}(K_{X_0(p)})}{2}$  for the image of Riemann's characteristic, which is some pre-image under duplication in  $J_0(p)$  of the image of some canonical divisor:  $\omega^0 = \iota_{P_0}(K_{X_0(p)})$  (cf. Theorem 4.3 above). The usual analytic norm of Riemann's theta function is

$$\|\theta(z)\|_{\text{an}} := \det(\Im(\tau))^{1/4} \exp(-\pi y \Im(\tau)^{-1} y) |\theta(z)| \quad (80)$$

for  $z = x + iy \in \mathbb{C}^g$  (cf. [45], (3.2.2)). That analytic metric is equal to the cubist one up to a multiplicative constant factor that can be computed with Moret-Bailly's "formule-clef", as we shall see in Lemmas 6.3 and 6.6 below.

Among the translates  $\Theta_D = t_D^* \Theta$ , for  $D \in J_0(p)(\mathbb{C})$ , of the above symmetric  $\Theta$ , the divisor  $\Theta_{\kappa} = t_{\kappa}^* \Theta = \sum_{i=1}^{g-1} \iota_{\infty}(X_0(p)_{\mathbb{Q}})$  defines an invertible sheaf  $L(\Theta_{\kappa})$  on  $J_0(p)$  over  $\mathbb{Q}$ . If  $\mathcal{N}_J^0$  denotes the neutral component of the Néron model of  $J$  over  $\mathbb{Z}$  and  $\mathcal{L}(\Theta_{\kappa})$  is the cubist extension of  $L(\Theta_{\kappa})$  to  $\mathcal{N}_J^0$  (cf. [44], Proposition II.1.2.2, as in Section 5.2 above), we know that  $H^0(\mathcal{N}_J^0, \mathcal{L}(\Theta_{\kappa}))$  is a (locally...) free  $\mathbb{Z}$ -module of rank 1, that the complex base-change  $H^0(J_0(p)(\mathbb{C}), L(\Theta_{\kappa, \mathbb{C}}))$  is similarly a complex line. So if  $s_{\theta}$  is a generator of the former space, whose image in the later we denote by  $s_{\theta, \mathbb{C}}$ , there is a nonzero complex number  $C_{\theta}$  such that

$$s_{\theta, \mathbb{C}}(z) = C_{\theta} \cdot \theta(z + \kappa). \quad (81)$$

Up to making some base-change from  $\mathbb{Z}$  to some  $\mathcal{O}_K$  we can now forget about  $\kappa$  and come back to the symmetric  $\Theta$ : we have  $[2N_0]^* \mathcal{L}(\Theta)|_{\mathcal{N}_J^0} \simeq \mathcal{L}(\Theta)^{\otimes 4N_0^2}$  on  $\mathcal{N}_{J, 2N_0}$  (cf. [50], Proposition 5.1) and we define a global section

$$s_{\mathcal{M}} := ([2N_0]^* t_{-\kappa}^*) s_{\theta} \in H^0(\mathcal{N}_{J, 2N_0}, [2N_0]^* \mathcal{L}(\Theta)_{\mathcal{O}_K}). \quad (82)$$

We will shortly show how to control the supremum of  $\|s_{\theta}\|_{\text{cub}}$ , therefore of  $\|s_{\mathcal{M}}\|_{\text{cub}}$ , on  $J_0(p)(\mathbb{C})$  (cf. Lemma 6.5). We shall moreover fix the morphism  $j_{\mathcal{M}}: \tilde{\mathcal{N}}_{J, 2N_0} \rightarrow \mathcal{J} \hookrightarrow \mathbb{P}_{\mathcal{O}_K}^n$  of (73) that is, the isomorphism  $j_{\mathcal{N}}^* \mathcal{O}_{\mathbb{P}^n}(1) \simeq \mathcal{M}$ , by mapping the canonical coordinates  $(x_i)_{0 \leq i \leq n}$  to sections  $(s_i)$  which will be translates of the above  $s_{\mathcal{M}}$  by  $4N_0^2$ -torsion points, as explained in Lemma 6.6 and its proof. (The sections  $s_i$ , as the  $x_i$ , are a priori defined on  $\tilde{\mathcal{N}}_{J, 4N_0^2}$  only, but one can consider their restrictions to  $\tilde{\mathcal{N}}_{J, 2N_0}$ .) This will allow us to also control the supremum of those  $s_i$ , relative to the cubist metrics, on the complex base change of our abelian varieties, as is required by the proof of arithmetic Bézout theorems.

We now start the technical preparation for the proof of Proposition 6.1, for which we need some Lemmas on the behavior of heights and degree under Segre maps, comparison between cubist and analytic metrics on theta functions, and estimates for all.

**Lemma 6.3** *Let  $(A, L)$  be a semistable, principally polarized abelian variety of dimension  $g$  over a number field  $K$ , with  $L$  a symmetric sheaf as in Proposition 6.1. For each embedding  $\sigma: K \hookrightarrow \mathbb{C}$  let  $\|\cdot\|_{\text{cub}, \sigma}$  denote the cubist metric on  $L$ , and  $\|\cdot\|_{\text{an}, \sigma}$  stand for the analytic metric on  $L_{\sigma} = \mathcal{L} \otimes_{\mathcal{O}_K, \sigma} \mathbb{C}$  defined in a similar fashion as in (80). Then*

$$\frac{1}{[K : \mathbb{Q}]} \sum_{\sigma: K \hookrightarrow \mathbb{C}} \log \|\cdot\|_{\text{cub}, \sigma} = \frac{1}{[K : \mathbb{Q}]} \sum_{\sigma: K \hookrightarrow \mathbb{C}} \log \|\cdot\|_{\text{an}, \sigma} + \left( \frac{1}{2} h_F(A) + \frac{g}{2} \log(2\pi) \right) \quad (83)$$

where  $h_F(A)$  is Faltings' (semistable) height of  $A$ .

**Proof** This directly follows from Moret-Bailly's formule-clef. At each complex place  $\sigma$ , both  $\|\cdot\|_{\text{cub}, \sigma}$  and  $\|\cdot\|_{\text{an}, \sigma}$  indeed have same (translation-invariant) curvature, so they differ from each

other by a constant multiplicative factor:  $\|\cdot\|_{\text{cub},\sigma} = e^{\phi_\sigma} \|\cdot\|_{\text{an},\sigma}$  (see [45] or [1], p. 150). Using a meromorphic section  $t_{\mathcal{L}}$  of the cubist extension  $\mathcal{L}$  of  $L$  to the neutral component  $\mathcal{N}_A^0$  of the Néron model of  $A$  (as before (72)) which does not vanish at the zero section  $0: \text{Spec}(\mathcal{O}_K) \rightarrow \mathcal{N}_A^0$ , the evaluation of the Arakelov degree of  $0^*(\mathcal{L})$  endowed with  $\|\cdot\|_{\text{cub}} := (\|\cdot\|_{\text{cub},\sigma})_{\{\sigma\}}$  gives

$$0 = \widehat{\deg}(0^*(\mathcal{L}, \|\cdot\|_{\text{cub}})) = \log(\text{card}(0^*(\mathcal{L})/\mathcal{O}_K \cdot t_{\mathcal{L}}(0))) - \sum_{\sigma} \log \|t_{\mathcal{L}}(0)\|_{\text{cub},\sigma}$$

so that

$$\sum_{\sigma} \phi_{\sigma} = \widehat{\deg}(0^*(\mathcal{L}, \|\cdot\|_{\text{an}})).$$

Now let  $\omega_A = 0^*(\bigwedge^g \Omega_{\mathcal{N}_A^0/\mathcal{O}_K})$  be the hermitian sheaf defining Faltings' semistable height of  $A$ , with metrics normalized such that

$$\|\alpha\|_{\sigma}^2 = \frac{i^{g^2}}{2g} \int_{A_{\sigma}(\mathbb{C})} \alpha \wedge \bar{\alpha}$$

at each  $\sigma$ . The formule-clef (Théorème 3.3 of [45]) writes

$$\widehat{\deg}(0^*(\mathcal{L}, \|\cdot\|_{\text{an}})) = \frac{1}{2} \widehat{\deg}(\omega_A) + [K : \mathbb{Q}] \frac{g}{2} \log(2\pi) = [K : \mathbb{Q}] \left( \frac{1}{2} h_F(A) + \frac{g}{2} \log(2\pi) \right).$$

□

**Lemma 6.4** *There is an infinite sequence  $(P_i)_{i \in \mathbb{N}}$  of points in  $X_0(p)(\overline{\mathbb{Q}})$  which are ordinary at all places dividing  $p$ , have everywhere integral  $j$ -invariant, and (with notations as in Theorem 4.3) normalized theta height of shape  $\hat{h}_{\Theta}(P_i - \infty + \frac{1}{2}\omega^0) = O(p^2)$ .*

**Proof** Let  $(\zeta_i)_{\mathbb{N}}$  be a infinite sequence of roots of unity of which none is congruent to some supersingular  $j$ -invariant in characteristic  $p$ , modulo any place of  $\mathbb{Q}$  above  $p$ . (As the supersingular  $j$ -invariants are quadratic over  $\mathbb{F}_p$ , one can for instance choose for the  $\zeta_i$  some primitive  $\ell_i$ -roots of unity, with  $\ell_i$  running through the set of primes (larger than  $p^2 - 1$ )). Lift each  $j$ -invariant equal to  $\zeta_i$  to some point  $P_i$  in  $X_0(p)(\overline{\mathbb{Q}})$ . By construction, this makes a sequence of everywhere ordinary points with  $j$ -height  $h_j(P_i)$  equal to 0. As for their (normalized) theta height, one sees from (65) that

$$\hat{h}_{\Theta}(P_i - \infty + \frac{1}{2}\omega^0) = \frac{1}{[K : \mathbb{Q}]} [P_i, \tilde{\omega}_{\Theta}]_{\mu_0} = \frac{-1}{[K(P_i) : \mathbb{Q}]} \sum_{\sigma: K(P_i) \hookrightarrow \mathbb{C}} g_{\mu_0}(\infty, \sigma(P_i)) + O(\log p) \quad (84)$$

as  $\tilde{\omega}_{\Theta} = (g \cdot \infty + \frac{1}{2}(\Phi_{\omega} + \Phi_{\Theta}) + c_{\Theta} X_{\infty})$ , and the contribution at finite places of  $[P_i, \infty]$  is 0, the other contributions being controlled by some  $O(\log(p))$  (see Remark 4.5 and references therein). It is therefore enough to bound the  $|g_{\mu_0}(\infty, \sigma(P_i))|$ .

Now  $|j(P_i)|_{\sigma} = 1$  for all  $\sigma: K(P_i) \hookrightarrow \mathbb{C}$ , so the corresponding elements  $\tau$  in the usual fundamental domain in Poincaré upper half-plane for  $X_0(p)$  or  $X(p)$ , and similarly the absolute values of  $q_{\tau} = e^{2i\pi\tau}$ , are absolutely bounded. (For an explicit estimate of this bound, one can check Corollary 2.2 of [5] - which for instance gives  $|q_{\tau}| \leq e^{2500}$ .) From this, running through the proof of Theorem 11.3.1 of [20] (and adapting it to the case of  $X_0(p)$  instead of  $X_1(pl)$ ), we deduce that the  $\sigma(P_i)$  do not belong to the open neighborhood, in the atlas of loc. cit., of the cusp  $\infty$  in  $X_0(p)(\mathbb{C})$ . Therefore Proposition 10.13 of [40] applies and gives, with notations of loc. cit.,

$$|g_{\mu_0}(\infty, \sigma(P_i))| = |g_{\mu_0}(\infty, \sigma(P_i)) - h_{\infty}(\sigma(P_i))| = O(p^2) \quad (85)$$

(see Theorem 11.3.1 of [20] and its proof). □



**Lemma 6.5** *Let  $s_\theta$  be the “theta function over  $\mathbb{Z}$ ”, that is, the global section introduced in (81). One has:*

$$\sup_{J_0(p)(\mathbb{C})} (\log \|s_\theta\|_{\text{cub}}) \leq O(p^3) \quad (86)$$

and even, under [4],

$$\sup_{J_0(p)(\mathbb{C})} (\log \|s_\theta\|_{\text{cub}}) \leq O(p \log p). \quad (87)$$

**Proof** Writing  $s_{\theta, \mathbb{C}}(z) = C_\theta \cdot \theta(z + \kappa)$  as in (81), we first endeavor to estimate  $|C_\theta|$ .

For  $D$  in  $J_0(p)(\mathbb{C})$ , written as the linear equivalence class of some divisor  $\sum_{i=1}^g (P_i - \infty)$  on  $X_0(p)(\mathbb{C})$ , we consider the embedding

$$\iota_{D+\kappa}: \begin{cases} X_0(p) & \hookrightarrow J_0(p) \\ P & \mapsto \text{cl}(P - \infty - \kappa - D) \end{cases}$$

and  $\kappa$  is Riemann’s characteristic (cf. (79) above and around). For such a  $D$  whose  $P_i$  are assumed to belong to  $X_0(p)(\overline{\mathbb{Q}})$ , we know from the proof of Theorem 4.3 ((57), (58)) that

$$h_\Theta(\iota_{D+\kappa}(P)) = \frac{1}{[K(P, D) : \mathbb{Q}]} [P, \sum_i P_i + \Phi_D + c_D X_\infty]_{\mu_0} \quad (88)$$

for some vertical divisor  $\Phi_D$  (which, by the way, will contribute at most  $O(\log p)$  to the height of points) and a contribution at the infinite vertical fiber given by some real number  $c_D$  (which explicitly is

$$c_D = \frac{1}{2[K(P, D) : \mathbb{Q}]} \left( -[\infty, \infty]_{\mu_0} - [\tilde{\kappa} + \tilde{D}, 2\infty + \tilde{\kappa} + \tilde{D}] \right).$$

(Here as in the proof of Theorem 4.3, the notation  $\tilde{\Delta}$  stands for the divisor  $\Delta$  modified by some vertical one (that with trivial  $C_\infty$ -components) so that the output belongs to the neutral component of the Picard functor.)

It is moreover well-known that there is an open subset in  $J_0(p)(\mathbb{C})$  (for the complex topology, and even the Zariski topology) in which all points  $D$  are such that

$$h^0(D + g \cdot \infty) := \dim_{\mathbb{C}} H^0(X_0(p)(\mathbb{C}), L(D + g \cdot \infty)_{\mathbb{C}}) = \dim_{\mathbb{C}} H^0(X_0(p)(\mathbb{C}), \iota_{D+\kappa}^* L(\Theta_{\mathbb{C}})) = 1,$$

so that  $\iota_{D+\kappa}^*(\Theta) = \sum_i P_i$  (the latter being an equality between effective divisors, not just a linear equivalence (cf. [24], pp. 336–340)). It follows that  $\text{div}(s_{\theta, \mathbb{C}}) \cap \iota_{D+\kappa}(X_0(p)(\mathbb{C})) = \cup_i \iota_{D+\kappa}(P_i)$ , or  $\text{div}(\iota_{D+\kappa}^* s_\theta) = \sum_i P_i$ . More precisely, extending basis to some ring of integers  $\mathcal{O}_K$  so that the  $P_i$  define sections of the minimal regular model of  $X_0(p)$  over  $\mathcal{O}_K$ , and making (if necessary) a further basis extension such that  $\mathcal{L}(\Theta)$  has a cubist extension on the whole Néron model of  $J_0(p)$  over  $\mathcal{O}_K$  (Proposition II.1.2.2 on p. 45 of [44], cf. paragraph 5.2 above), one sees that  $s_\theta$  defines a meromorphic section of  $\mathcal{L}(\Theta)_{\mathcal{O}_K}$  and  $\text{div}(\iota_{D+\kappa}^*(s_\theta))$  has indeed to be  $(\sum_i P_i + \Phi_D)$  on  $X_0(p)_{\mathcal{O}_K}^{\text{smooth}}$ .

Now consider only points  $P_i$  occurring in Lemma 6.4 (which, in particular, are ordinary and have integral  $j$ -invariants): this we can of course do because those make a Zariski-dense subset of  $X_0(p)(\overline{\mathbb{Q}})$  (and the onto-ness of the map  $X_0(p)^{(g)} \xrightarrow{\iota_\infty^g} J_0(p)$ ). By Lemma 6.4 and (64),

$$h_\Theta(\iota_{D+\kappa}(\infty)) = \left\| -\left(\sum_i P_i - \infty\right) + \frac{1}{2}\omega^0 \right\|_\Theta^2 = O(p^4). \quad (89)$$

On the other hand, for some of those choices of  $(P_i)_{1 \leq i \leq g}$ , our  $\mathbb{Z}$ -theta function  $s_\theta$  does not vanish at  $\iota_{D+\kappa}(\infty)(\mathbb{Q})$ , so  $h_\Theta(\iota_{D+\kappa}(\infty))$  can also be computed as the Arakelov degree:

$$h_\Theta(\iota_{D+\kappa}(\infty)) = \widehat{\deg}(\infty^* \iota_{D+\kappa}^*(\mathcal{L}(\Theta)))$$

(on  $\text{Spec}(\overline{\mathbb{Z}})$ ). By integrality of the  $P_i$  there is no finite contribution to this degree (except vertical components; cf. (88)), that is, if  $K_D$  is a sufficiently large field, over which  $D$  is defined, then for a set of relevant elements  $(z_\sigma)_{\sigma: K_D \hookrightarrow \overline{\mathbb{Q}}}$  (which lift  $\sigma(D)$ ) in the complex tangent space of  $J_0(p)$  to 0 one has:

$$\begin{aligned} h_\Theta(\iota_{D+\kappa}(\infty)) &= \widehat{\deg}(0_{\mathcal{J}_0(p)}^*(t_{D+\kappa}^* \mathcal{L}(\Theta))) = \widehat{\deg}(0_{\mathcal{J}_0(p)}^*(t_D^* \mathcal{L}(\Theta_\kappa))) \\ &= \left(-\frac{1}{[K_D : \mathbb{Q}]}\right) \sum_{K_D \xrightarrow{\sigma} \mathbb{C}} \log \|s_\theta(z_\sigma)\|_{\text{cub}} + O(\log p). \end{aligned}$$

whence as  $s_{\theta, \mathbb{C}}(z) = C_\theta \cdot \theta(z + \kappa)$ :

$$\log |C_\theta| = -h_\Theta(\iota_{D+\kappa}(\infty)) - \frac{1}{[K_D(\kappa) : \mathbb{Q}]} \sum_{K_D(\kappa) \xrightarrow{\sigma} \mathbb{C}} \log \|\theta((z + \kappa)_\sigma)\|_{\text{cub}} + O(\log p). \quad (90)$$

By Lemma 6.3,

$$\log \|\theta(z)\|_{\text{cub}} = \log \|\theta(z)\|_{\text{an}} + \frac{1}{2} \left( h_F(J_0(p)) + \frac{g}{2} \log(2\pi) \right)$$

and, with notations of [22], paragraph 8, one has  $|F(z)| = 2^{g/4} \|\theta(z)\|_{\text{an}}$ . Indeed there is a constant  $A \in \mathbb{R}_+^*$  such that  $|F(z)| = A \cdot \|\theta(z)\|_{\text{an}}$  (cf. end of proof of Lemma 8.3 of [22]),  $\int_{J_0(p)(\mathbb{C})} |F|^2 d\nu = 1$  (where  $d\nu$  is the probability Haar measure on  $J_0(p)(\mathbb{C})$ ; cf. [22], Lemma 8.2 (1)), and  $\int_{J_0(p)(\mathbb{C})} \|\theta(z)\|_{\text{an}}^2 d\nu = 2^{-g/2}$  (see e.g [45], (3.2.1)). Therefore Lemma 8.3 of [22] gives:

$$-\frac{1}{[K_D(\kappa) : \mathbb{Q}]} \sum_{K_D(\kappa) \xrightarrow{\sigma} \mathbb{C}} \log \|\theta((z + \kappa)_\sigma)\|_{\text{cub}} \leq h_\Theta(\iota_{D+\kappa}(\infty)) + \frac{g}{4} \log 2$$

from which (90) implies:

$$\log |C_\theta| \leq \frac{g}{4} \log 2 + O(\log p).$$

Given this upper bound for  $|C_\theta|$ , we can now go the other way round to derive an upper bound for  $\|s_\theta\|_{\text{cub}} = C_\theta \cdot \|\theta(z + \kappa)\|_{\text{cub}}$ , by using general estimates for analytic theta functions. One can for instance invoke works of Igusa and Edixhoven-de Jong, who assert that for any principally polarized complex abelian variety whose complex invariant  $\tau$  is chosen within Siegel's fundamental domain  $F_g$  one has (recall (80)):

$$\frac{1}{\det(\Im(\tau))^{1/4}} \|\theta(z)\|_{\text{an}} = \exp(-\pi y \Im(\tau)^{-1} y) |\theta(z)| \leq 2^{3g^3+5g},$$

cf. [20], pp. 231-232. This upper bound has in fact just been announced to be amenable to the much better  $g^{g/2}$  by P. Autissier in [4].

We finally deal with the factor  $\det(\Im(\tau))^{1/4}$ . Lemma 11.2.2 of [20] is again completely general so we use it as it stands:

$$\det(\Im(z))^{1/2} \leq \frac{(2g)! V_{2g}}{2^g V_g} \prod_{g+1 \leq i \leq 2g} \lambda_i$$

where for any  $k$  we write  $V_k$  for the volume of the unit ball in  $\mathbb{R}^k$  endowed with its standard Euclidean structure, and the  $\lambda_r$  are the successive minima, relative to the Riemann form, of the lattice  $\Lambda = \mathbb{Z}^g + \tau \cdot \mathbb{Z}^g$ . To bound the  $\lambda_i$  we need to invoke an avatar of loc. cit., Lemma 11.2.3. But the very same proof shows that for any integer  $N$ , the group  $\Gamma_0(N)$  has a set of generators having entries of absolute value less or equal to the very same bound  $N^6/4$ . (That term could be improved, but it has an invisible impact on the final bounds so we here content ourselves with it.) We can therefore rewrite the proof of Lemma 11.2.4 verbatim. This gives that  $\Lambda$  is generated by elements having naive hermitian norm  $\|x\|_E^2$  less or equal to  $gp^{46}$ . Finally, in our case the Gram matrix is diagonal (no  $2 \times 2$ -blocks, at the difference of Lemma 11.1.4 of loc. cit.)

so Lemma 11.2.5 a fortiori holds: if  $\|\cdot\|_P$  denotes the hermitian product on  $\mathbb{C}^g$  induced by the polarization,  $\|\cdot\|_P^2 \leq \frac{e^{4\pi}}{\pi} \|\cdot\|_E^2$ . This allows to conclude as in p. 228 of [20]:

$$\left(\prod_{i=g+1}^{2g} \lambda_i\right)^2 \leq \left(\frac{e^{4\pi}}{\pi} gp^{46}\right)^g$$

so that finally

$$\log(\det(\Im(\tau))) = O(p \log p).$$

As we know that  $h_F(J_0(p)) = O(p \log p)$  again (Théorème 1.2 of [60]), comparing analytic and cubist metric thanks to (83) and putting everything together we finally obtain:

$$\log \|s_{\theta, \mathbb{C}}(z)\|_{\text{cub}} = \log \|C_{\vartheta} \cdot \theta(z + \kappa)\|_{\text{cub}} \leq O(p^3)$$

for all  $z$  in  $J_0(p)(\mathbb{C})$ , using Igusa-Edixhoven-de Jong estimates above, and even

$$\sup_{J_0(p)(\mathbb{C})} \log \|s_{\theta, \mathbb{C}}\|_{\text{cub}} = O(p \log p)$$

under Autissier's result [4]. □

**Lemma 6.6** *Assume the same hypothesis and notations as in Definition 5.3 and Lemma 6.5. After possibly making some finite extension of basis one can pick a set  $\mathcal{S}$  in  $H^0(\mathcal{N}_{J, 2N_0}, \mathcal{L}(\Theta)^{\otimes 4N_0^2})$  of  $(2N_0)^{2g}$  global sections  $(s_i)_{1 \leq i \leq (2N_0)^{2g}}$  which span  $\mathcal{L}(\Theta)^{\otimes 4N_0^2}$  on  $\mathcal{N}_{J, 2N_0}$  and verify*

$$\sup_{J_0(p)_{\sigma}(\mathbb{C})} (\log \|s_i\|_{\text{cub}}) \leq O(p^3) \tag{91}$$

and even, under Autissier's [4]:

$$\sup_{J_0(p)_{\sigma}(\mathbb{C})} (\log \|s_i\|_{\text{cub}}) \leq O(p \log p). \tag{92}$$

**Proof** Up to making an extension of basis, we can assume  $L(\Theta)^{\otimes 4N_0^2}$  (respectively,  $[2N_0]^* L(\Theta)$ ) has a cubist extension  $\mathcal{L}(\Theta)^{\otimes 4N_0^2}$  (respectively,  $[2N_0]^* \mathcal{L}(\Theta)$ ) on  $\mathcal{N}_{J, 4N_0^2}$ . One then knows that, as  $\Theta$  is symmetric, there is an isomorphism  $[2N_0]^* \mathcal{L}(\Theta) \rightarrow \mathcal{L}(\Theta)^{\otimes 4N_0^2}$  which actually is an isometry ([50], Proposition 5.1), by which we identify those two objects from now on. On the other hand, by our construction, every element  $x$  of  $J_0(p)[4N_0^2](\overline{\mathbb{Q}}) = J[4N_0^2](K)$  defines a section  $\tilde{x}$  in  $\mathcal{N}_{J, 4N_0^2}(\text{Spec}(\mathcal{O}_K))$  (cf. paragraph 5.2 above). Letting  $t_{\tilde{x}}$  denote the translation by  $\tilde{x}$  on  $\mathcal{N}_{J, 4N_0^2}$  we have

$$t_{\tilde{x}}^* \mathcal{L}(\Theta)^{\otimes 4N_0^2} \simeq \mathcal{L}(\Theta)^{\otimes 4N_0^2}. \tag{93}$$

(This is indeed true over  $\mathbb{C}$  by Lemma 2.4.7.c) of [7], hence over  $K$ , then over  $\text{Spec}(\mathcal{O}_K)$  by uniqueness of cubist extensions.) The interpretation as Néron-Tate heights shows that as  $\mathcal{L}(\Theta)$  is endowed with its cubist metric, this isomorphism even is an isometry. So if  $s_{\mathcal{M}}$  is the section defined in (82), we define a set  $(s_i)_{1 \leq i \leq (2N_0)^{2g}}$  in  $H^0(\mathcal{N}_{J, 2N_0}, [2N_0]^* \mathcal{L}(\Theta))$  made of  $(2N_0)^{2g}$  elements of shape

$$s_i := t_{\tilde{x}_i}^* s_{\mathcal{M}} \tag{94}$$

for  $\tilde{x}_i$  running through a set of representative, in  $J_0(p)[4N_0^2](K)$ , of  $(J_0(p)[4N_0^2]/J_0(p)[2N_0])$ . Note that one can explicitly lift  $s_{\mathcal{M}}$  as

$$s_{\mathcal{M}, \mathbb{C}}(z) = C_{\vartheta} \cdot \theta(2N_0 \cdot z) \tag{95}$$

on the complex tangent space to 0 of  $J_0(p)(\mathbb{C})$  (and the  $s_{i,\mathbb{C}}$  are constant multiple of the basis denoted by  $h_{\vec{a},\vec{b}}(\vec{z})$  in [47], Proposition II.1.3.iii) on p. 124<sup>7</sup>). From which Lemma 6.5 gives (91) and (92).

By the theory of  $\theta$ -functions (cf. [50], Proposition 2.5 and its proof, [46] and [44], Chapitre VI) the  $s_i$  make a basis of global sections (which span  $\mathcal{L}(\Theta)^{\otimes 4N_0^2}$  everywhere).  $\square$

**Remark 6.7** For places above  $p$ , one checks directly that the global sections  $(s_i)$  of Lemma 6.6 have no common zeroes on  $\mathcal{N}_{J,2N_0}$ , so that those  $s_i$  can indeed be used to build the projective embeddings of (73). Indeed, the divisor of  $s_\theta$  (restricted to  $\mathcal{N}_{J,2N_0}$ ) is the image of  $\mathcal{X}_0(p)^{(g-1)}$  in  $\mathcal{N}_{J,2N_0}$ , which is not equal to all points of  $\mathcal{N}_{J,2N_0}(\mathbb{F}_p)$  (a form of  $\mathbb{G}_m^g$ ). So at any point of  $\mathcal{N}_{J,2N_0}(\mathbb{F}_p)$ , at least one of the  $s_i$  has not to vanish.

**Lemma 6.8** *Let  $V$  and  $W$  be two  $K$ -subvarieties of a polarized abelian variety  $(A, M)$  over a number field  $K$ , of dimension  $d_V$  and  $d_W$  respectively. Assume the polarized arithmetic scheme  $(\mathcal{A}, \mathcal{M})$ , with  $\mathcal{M}$  an hermitian sheaf on  $\mathcal{A}$ , is a model for  $(A, M)$  over  $\text{Spec}(\mathcal{O}_K)$ . Then*

$$\deg_{M^{\boxtimes 2}}(V \times W) = \binom{d_V + d_W}{d_V} \deg_M(V) \deg_M(W) \quad (96)$$

and

$$\begin{aligned} h_{\mathcal{M}^{\boxtimes 2}}(\mathcal{V} \times \mathcal{W}) &= \binom{d_V + d_W + 1}{d_V} \deg_M(V) h_{\mathcal{M}}(\mathcal{W}) \\ &\quad + \binom{d_V + d_W + 1}{d_W} \deg_M(W) h_{\mathcal{M}}(\mathcal{V}). \end{aligned} \quad (97)$$

Moreover, (97) remains true with Néron-Tate normalized heights  $\hat{h}_M$  and  $\hat{h}_{M^{\boxtimes 2}}$  in place of  $h_{\mathcal{M}}$  and  $h_{\mathcal{M}^{\boxtimes 2}}$  respectively.

**Remark 6.9** The two equalities of Lemma 6.8 give

$$\frac{h_{\mathcal{M}}(\mathcal{V} \times \mathcal{W})}{\dim(\mathcal{V} \times \mathcal{W}) \deg_M(V \times W)} = \frac{h_{\mathcal{M}}(\mathcal{V})}{\dim(\mathcal{V}) \deg_M(V)} + \frac{h_{\mathcal{M}}(\mathcal{W})}{\dim(\mathcal{W}) \deg_M(W)}$$

whose three terms are the “normalized” height, in the sense of [1], Proposition-Définition 2.1 (up to a common multiplication by  $[K : \mathbb{Q}]$ ), of the respective objects. This of course fits with the obvious naive interpretation in terms of essential minima (cf. the beginning of Section 5 in the present paper).

**Proof (of Lemma 6.8).** As for (96), one can realize it is elementary, or refer to Lemme 2.2 of [55], or proceed as follows. With notations of [10], (2.3.18) and (2.3.19), using Proposition 3.2.1, (iii) of loc. cit., and noticing  $c_1(\mathcal{M}^{\boxtimes 2}) = c_1(\mathcal{M}) \times \mathbf{1} + \mathbf{1} \times c_1(\mathcal{M})$  one computes:

$$\begin{aligned} \deg_{\mathcal{M}^{\boxtimes 2}}(V \times W) &= \deg_{\mathbb{Q}}(c_1(\mathcal{M}^{\boxtimes 2})^{d_V + d_W} | (V \times W)) \\ &= \deg_{\mathbb{Q}}\left(\sum_{k=0}^{d_V + d_W} \binom{d_V + d_W}{k} c_1(\mathcal{M})^k \times c_1(\mathcal{M})^{d_V + d_W - k} | V \times W\right) \\ &= \sum_{k=0}^{d_V + d_W} \binom{d_V + d_W}{k} \deg_{\mathbb{Q}}(c_1(\mathcal{M})^k \times c_1(\mathcal{M})^{d_V + d_W - k} | V \times W) \\ &= \sum_{k=0}^{d_V + d_W} \binom{d_V + d_W}{k} \deg_{\mathbb{Q}}(c_1(\mathcal{M})^k | V) \deg_{\mathbb{Q}}(c_1(\mathcal{M})^{d_V + d_W - k} | W). \end{aligned}$$

---

<sup>7</sup>where it seems by the way that the expression “ $h_{\vec{a},\vec{b}}(\vec{z}) = \vartheta[\frac{\vec{a}/k}{\vec{b}/k}](\ell \cdot \vec{z}, \Omega)$ ” should read “ $\dots = \vartheta[\frac{\vec{a}/k}{\vec{b}/k}](k \cdot \vec{z}, \Omega)$ ” (notations of loc. cit.).

The only nonzero term in the last sum occurs for  $k = d_V$  and we obtain

$$\deg_{\mathcal{M}^{\boxtimes 2}}(V \times W) = \binom{d_V + d_W}{d_V} \deg_{\mathbb{Q}}(c_1(\mathcal{M})^{d_V} | V) \deg_{\mathbb{Q}}(c_1(\mathcal{M})^{d_W} | W)$$

which is (96).

An analogous computation, using [10], (2.3.19), can be led for the height:

$$\begin{aligned} h_{\mathcal{M}^{\boxtimes 2}}(\mathcal{V} \times \mathcal{W}) &= \widehat{\deg}(c_1(\mathcal{M}^{\boxtimes 2})^{d_V + d_W + 1} | \mathcal{V} \times \mathcal{W}) \\ &= \widehat{\deg}\left(\sum_{k=0}^{d_V + d_W + 1} \binom{d_V + d_W + 1}{k} c_1(\mathcal{M})^k \times c_1(\mathcal{M})^{d_V + d_W + 1 - k} | \mathcal{V} \times \mathcal{W}\right) \\ &= \binom{d_V + d_W + 1}{d_V} \deg_{\mathbb{Q}}(c_1(\mathcal{M})^{d_V} | \mathcal{V}) \widehat{\deg}(c_1(\mathcal{M})^{d_W + 1} | \mathcal{W}) + \\ &\quad \binom{d_V + d_W + 1}{d_W} \widehat{\deg}(c_1(\mathcal{M})^{d_V + 1} | \mathcal{V}) \deg_{\mathbb{Q}}(c_1(\mathcal{M})^{d_W} | \mathcal{W}) \\ &= \binom{d_V + d_W + 1}{d_V} \deg_{\mathcal{M}}(\mathcal{V}) h_{\mathcal{M}}(\mathcal{W}) + \binom{d_V + d_W + 1}{d_W} \deg_{\mathcal{M}}(\mathcal{W}) h_{\mathcal{M}}(\mathcal{V}). \end{aligned}$$

From here the deduction of the statement relative to Néron-Tate heights is formal.  $\square$

**Proof (of Proposition 6.1).** Construction (77) gives us a  $\mathbb{Q}$ -embedding  $V \times W \xhookrightarrow{\iota} \mathbb{P}^{n^2 + 2n}$  via a Segre map. We set

$$s_{\underline{i}, \underline{j}} := \iota^*(z_{\underline{i}, \underline{j}} - z_{\underline{j}, \underline{i}})$$

for all  $(\underline{i}, \underline{j})$ , and denote by  $\mathcal{O}_{N_0}$  the ambient line bundle  $\iota^*(\mathcal{O}_{\mathbb{P}^{n^2 + 2n}}(1)) = \mathcal{M}^{\boxtimes 2}$  as before (78). Set also  $\bar{\mathcal{O}}_{N_0} := \mathcal{O}_{N_0} \otimes \mathbb{Q}$ . We intersect  $\iota(V \times W)$  with one of the  $\text{div}(z_{\underline{i}_0, \underline{j}_0} - z_{\underline{j}_0, \underline{i}_0})_{\mathbb{Q}}$  such that the two cycles meet properly: define

$$J_1 = \text{div}(s_{\underline{i}_0, \underline{j}_0}) \cap (V \times W)$$

in the generic fiber  $(J_0(p) \times J_0(p))_{\mathbb{Q}}$ . As  $\text{div}(z_{\underline{i}_0, \underline{j}_0} - z_{\underline{j}_0, \underline{i}_0})$  is a projective hyperplane we have by definition

$$\deg_{\mathcal{O}_{N_0}}(J_1) = \deg_{\mathcal{O}_{N_0}}(V \times W).$$

For the same linearity reason, a similar statement is true for heights. Let indeed  $\mathcal{V}$  and  $\mathcal{W}$  denote the schematic closure in  $\mathcal{J}$  of  $V$  and  $W$  respectively, and  $\mathcal{J}_1$  the schematic closure of  $J_1$  in  $\mathcal{J} \times \mathcal{J}$ , which satisfies

$$h_{\mathcal{O}_{N_0}}(\mathcal{J}_1) \leq h_{\mathcal{O}_{N_0}}(\text{div}(s_{\underline{i}_0, \underline{j}_0}) \cap (\mathcal{V} \times \mathcal{W})).$$

Proposition 3.2.1 (iv) of [10] gives, with notations of loc. cit., that:

$$\begin{aligned} h_{\mathcal{O}_{N_0}}(\text{div}(s_{\underline{i}_0, \underline{j}_0}) \cap (\mathcal{V} \times \mathcal{W})) &= h_{\mathcal{O}_{N_0}}(\mathcal{V} \times \mathcal{W}) \\ &\quad + \frac{1}{[K : \mathbb{Q}]} \sum_{\sigma : K \hookrightarrow \mathbb{C}} \int_{(V \times W)_{\sigma}(\mathbb{C})} \log \|s_{\underline{i}_0, \underline{j}_0}\|_{c_1(\mathcal{O}_{N_0})^{d_V + d_W}} \end{aligned} \quad (98)$$

where  $\|\cdot\| = \|\cdot\|_{\text{cub}}$  shall denote the cubist metric (or the metric induced by the cubist metric on products or powers of relevant sheaves). To estimate the last integral, we note that at any point of  $(V \times W)_{\sigma}(\mathbb{C})$  and for any  $(\underline{i}, \underline{j})$ ,

$$\begin{aligned} \|s_{\underline{i}, \underline{j}}\| &= \|z_{\underline{i}, \underline{j}} - z_{\underline{j}, \underline{i}}\|_{\mathcal{M}^{\boxtimes 2}} \leq \|z_{\underline{i}, \underline{j}}\|_{\mathcal{M}^{\boxtimes 2}} + \|z_{\underline{j}, \underline{i}}\|_{\mathcal{M}^{\boxtimes 2}} \\ &\leq \|x_{\underline{i}}\|_{\mathcal{M}} \|y_{\underline{j}}\|_{\mathcal{M}} + \|x_{\underline{j}}\|_{\mathcal{M}} \|y_{\underline{i}}\|_{\mathcal{M}} \leq 2(\sup_{\underline{i}} \|x_{\underline{i}}\|_{\mathcal{M}})^2 \\ &\leq \exp(2 \log(\sup \|s_i\|_{\text{cub}}) + \log 2) \end{aligned}$$

with notations of Lemma 6.6 above. Setting  $M_{\mathcal{J},\mathcal{M}} = \log(\sup \|s_i\|_{\text{cub}})$  we obtain

$$h_{\mathcal{O}_{N_0}}(\mathcal{I}_1) \leq h_{\mathcal{O}_{N_0}}(\mathcal{V} \times \mathcal{W}) + (2M_{\mathcal{J},\mathcal{M}} + \log 2) \deg_{\mathcal{O}_{N_0}}(V \times W).$$

Call  $I_1$  one of the reduced irreducible components of  $J_1$  containing the point  $\iota(\Delta(P))$  of  $V \cap W$  considered in the statement of Proposition 6.1, and let  $\mathcal{I}_1$  denote its Zariski closure in  $\mathcal{J}$ . It has  $\mathcal{O}_{N_0}$ -height (and degree) less or equal to those of  $\mathcal{J}_1$ , so that again

$$h_{\mathcal{O}_{N_0}}(\mathcal{I}_1) \leq h_{\mathcal{O}_{N_0}}(\mathcal{V} \times \mathcal{W}) + (2M_{\mathcal{J},\mathcal{M}} + \log 2) \deg_{\mathcal{O}_{N_0}}(V \times W)$$

and we can iterate the process with  $I_1$  in place of  $V \times W$ : we obtain some  $J_2, \mathcal{J}_2, I_2, \mathcal{I}_2$  such that

$$\begin{aligned} h_{\mathcal{O}_{N_0}}(\mathcal{I}_2) &\leq h_{\mathcal{O}_{N_0}}(\mathcal{I}_1) + (2M_{\mathcal{J},\mathcal{M}} + \log 2) \deg_{\mathcal{O}_{N_0}}(I_1) \\ &\leq h_{\mathcal{O}_{N_0}}(\mathcal{V} \times \mathcal{W}) + 2(2M_{\mathcal{J},\mathcal{M}} + \log 2) \deg_{\mathcal{O}_{N_0}}(V \times W). \end{aligned}$$

(The only obstruction to this step is if all the  $s_{k,l}$  vanish on  $I_1$ , which implies it is contained in the diagonal of  $J_0(p) \times J_0(p)$  - so that  $I_1 = \iota(\Delta(P))$  by construction and we are already done.) Processing, one builds a sequence  $(\mathcal{I}_k)$  of integral closed subschemes of  $\mathcal{J} \times \mathcal{J}$ , with decreasing dimension, such that the last step gives

$$h_{\mathcal{O}_{N_0}}(\mathcal{I}_{d_V+d_W}) \leq h_{\mathcal{O}_{N_0}}(\mathcal{V} \times \mathcal{W}) + (d_V + d_W)(2M_{\mathcal{J},\mathcal{M}} + \log 2) \deg_{\mathcal{O}_{N_0}}(V \times W).$$

Now  $h_{\mathcal{O}_{N_0}}(\mathcal{I}_{d_V+d_W}) = h_{\mathcal{O}_{N_0}}(\Delta(P, P)) = h_{\mathcal{M}^{\otimes 2}}(P) = h_{\mathcal{L}^{\otimes 8N_0^2}}(P) = 8N_0^2 \hat{h}_L(P)$  (Néron-Tate normalized theta height). (Note that the equality  $h_{\mathcal{L}^{\otimes 4N_0^2}}(P) = h_{\mathcal{M}}(P)$  in the above line comes from the fact that  $\mathcal{M}$ , a priori a subsheaf of  $\mathcal{L}^{\otimes 4N_0^2}$ , actually coincides with it along the section  $P(\overline{\mathbb{Z}})$ , as the section  $s_i$  of Lemma 6.6 span  $\mathcal{L}^{\otimes 4N_0^2}$  along  $P$  (cf. Remark 6.7 and Proposition 5.4)). Using Lemma 6.8 and Corollary 5.5 we therefore obtain

$$\begin{aligned} 8N_0^2 \hat{h}_L(P) &\leq (4N_0^2)^{d_V+d_W+1} \left[ (d_W + 1) \binom{d_V + d_W + 1}{d_V} \hat{h}_{\Theta}(W) \deg_{\Theta}(V) \right. \\ &\quad \left. + (d_V + 1) \binom{d_V + d_W + 1}{d_W} \hat{h}_{\Theta}(V) \deg_{\Theta}(W) \right] \\ &\quad + (d_V + d_W)(2M_{\mathcal{J},\mathcal{M}} + \log 2) \binom{d_V + d_W}{d_V} (4N_0^2)^{d_V+d_W} \deg_{\Theta}(V) \deg_{\Theta}(W). \end{aligned}$$

From here, Lemma 6.6 allows to conclude the proof of Proposition 6.1.  $\square$

That arithmetic Bézout theorem will be our principal tool in the sequel.

## 7 Height bounds for quadratic points on $X_0(p^2)$

**Proposition 7.1** *Let  $\iota: X \hookrightarrow J$  be some Albanese map from a curve (of positive genus) over some field  $K$  to its jacobian  $J$ . Let  $\pi: J \rightarrow A$  be some quotient of  $J$ , with  $\dim(A) > 1$ , and  $X'$  be the normalization of the image  $\pi \circ \iota(X)$  of  $X$  in  $A$ . Then the map  $\pi': X \rightarrow X'$  induced by  $\pi \circ \iota$  verifies*

$$\deg(\pi') \leq \frac{\dim(J) - 1}{\dim(A) - 1}.$$

**Proof** The map  $\pi \circ \iota$  induces an inclusion of function fields which defines the map  $\pi': X \rightarrow X'$ . If  $J'$  is the jacobian of  $X'$ , Albanese functoriality says that  $\pi$  factorizes through surjective morphisms  $J \rightarrow J' \rightarrow A$ . Hurwitz formula says:

$$\deg(\pi') = \frac{\dim(J) - 1 - \frac{1}{2} \deg R}{\dim(J') - 1}$$

for  $R$  the ramification divisor of  $\pi'$ . Whence the result.  $\square$

**Lemma 7.2** Assume  $p \geq 23$ ,  $p \neq 37$ . Let  $X := X_0(p)$  and  $\pi_e: J_0(p) \rightarrow J_e$  the projection. Let

$$\iota_{P_0}: \begin{cases} X_0(p) & \hookrightarrow J_0(p) \\ P & \mapsto \text{cl}(P - P_0) \end{cases}$$

for some  $P_0$  in  $X_0(p)(\overline{\mathbb{Q}})$  such that  $w_p(P_0) = P_0$  (there are roughly  $\sqrt{p}$  such points, cf. Proposition 3.1 of [27]), and set  $\varphi_e := \pi_e \circ \iota_{P_0}$ . Then:

- the equality

$$\varphi_e(X_0(p)) = a - \varphi_e(X_0(p)) = a + w_p \varphi_e(X_0(p)) = a + \varphi_e(X_0(p)), \quad (99)$$

for a in  $J_e(\mathbb{Q})$  some (necessarily torsion) point, implies  $a = 0$ ;

- If  $d$  is the degree of the map  $X_0(p) \rightarrow \widetilde{\varphi_e(X_0(p))}$  to the normalization of  $\varphi_e(X_0(p))$ , then  $d$  is either 1, 3 or 4;
- Assuming moreover Brumer's conjecture (cf. (18), and also (19)), then equality (99) implies  $d = 1$  for large enough  $p$ .

**Proof** Notice first that, by our choice of  $P_0$  (whence  $\iota$ ), and because  $J_e$  belongs to the  $w_p$ -minus part of  $J_0(p)$ , one has:

$$\varphi_e(w_p(P)) = w_p(\varphi_e(P)) = -\varphi_e(P)$$

for all  $P \in X_0(p)(\mathbb{C})$ . So let  $n$  be the order of  $a$ . We remark that the degree  $d$  cannot be equal to 2, as otherwise the extension of fraction fields  $K(X_0(p))/K(\varphi_e(X_0(p)))$  would be Galois and  $X_0(p)$  would possess an involution different from  $w_p$ , which is not by Ogg's theorem (see [49] or even [32]). If  $d = 1$ , the same reason that  $\text{Aut}(X_0(p)) = \langle w_p \rangle$  implies that  $n = 1$ . Let now  $X'$  be the normalization of the quotient of  $\varphi_e(X_0(p))$  by the automorphism  $P \mapsto P + a$ , that is, the image of  $\varphi_e(X_0(p))$  by the quotient morphism  $J_e \rightarrow J_e/\langle a \rangle$ . Let  $\pi$  be the composed map  $J_0(p) \xrightarrow{\varphi_e} J_e \rightarrow J_e/\langle a \rangle$ . The degree of  $X_0(p) \rightarrow X'$  is  $d \cdot n$  and Proposition 7.1 together with the left part of inequalities (20) implies:

$$d \cdot n \leq \frac{g-1}{(\frac{1}{4}-\varepsilon)g-1} \leq 4 + \varepsilon$$

for large enough  $p$ . This shows that if  $d = 3$  or 4 one still has  $a = 0$ , whence the Proposition's first two statements. Under Brumer's conjecture,  $d \cdot n \leq 2 + \varepsilon$ , so that  $d = 1$  and  $a = 0$  by previous arguments.  $\square$

**Remark 7.3** Replace, in Lemma 7.2, the map  $X_0(p) \rightarrow J_e$  by  $X_0(p) \rightarrow J_0(p)^-$  (by which the former factorizes, by the way). The above proof shows that the map  $X_0(p) \rightarrow \varphi(X_0(p))$  is of generic degree 1 (independently on any conjecture), but of course it needs not being injective on points: a finite number of points can be mapped together to singular points on  $\varphi(X_0(p))$ . In our case one checks those are among the Heegner points  $P$  such that  $P = w_p(P)$  (for which we again refer to Proposition 3.1 of [27]). Indeed, the endomorphism of  $J_0(p)$  defined by multiplication by  $(1 - w_p)$  factorizes through  $\varphi$  (and  $\cdot(1 - w_p)$  is the map considered in (2) and around, inducing multiplication by 2 on tangent spaces). Therefore, if  $P$  maps to a multiple point of  $\varphi(X_0(p))$ , it also maps to a multiple point of  $(1 - w_p) \circ \iota(X_0(p))$ . Now, assuming  $X_0(p)$  has gonality larger than 2 (which is true as soon as  $p > 71$ , cf. [48], Theorem 2), the equality  $\text{cl}((1 - w_p)P) = \text{cl}((1 - w_p)P')$  in  $J_0(p)$ , for some  $P'$  on  $X_0(p)$  different from  $P$ , implies  $P = w_p P$ ,  $P' = w_p P'$ . That is,  $P$  (and  $P'$ ) are Heegner points.

**Lemma 7.4** Suppose  $P$  belongs to  $X_0(p^2)(K)$  for some quadratic number field  $K$ , and  $P$  is not a complex multiplication point. Then for one of the two natural degeneracy morphisms  $\pi$  from  $X_0(p^2)$  to  $X_0(p)$ , the point  $Q := \pi(P)$  in  $X_0(p)(K)$  does not define a  $\mathbb{Q}$ -valued point of the quotient curve  $X_0^+(p) := X_0(p)/w_p$ .

**Proof** Using the modular interpretation, we write  $P = (E, C_{p^2})$  for  $E$  an elliptic curve over  $K$  and  $C_{p^2}$  a cyclic  $K$ -isogeny of degree  $p^2$ . One deduces from it the two points  $Q_1 := (E, p \cdot C_{p^2})$  and  $Q_2 := (E/p \cdot C_{p^2}, C_{p^2} \bmod p \cdot C_{p^2})$  in  $X_0(p)(K)$ . Assume both  $Q_1$  and  $Q_2$  do define elements of  $X_0^+(p)(\mathbb{Q})$ . If  $\sigma$  denotes a generator of  $\text{Gal}(K/\mathbb{Q})$ , the above means that

$$w_p(Q_1) = (E/p \cdot C_{p^2}, E[p] \bmod p \cdot C_{p^2}) \simeq \sigma(Q_1)$$

and

$$w_p(Q_2) = (E/C_{p^2}, E[p] + C_{p^2} \bmod C_{p^2}) \simeq \sigma(Q_2).$$

Therefore  $E \simeq {}^\sigma(E/p \cdot C_{p^2}) \simeq E/C_{p^2}$ , so that  $E$  has complex multiplication.  $\square$

We can now conclude with the main result of this paper.

**Theorem 7.5** *Assume Brumer's conjecture (18) holds. There is an integer  $C$  such that, if  $P$  is a quadratic point of  $X_0(p^2)$  (that is:  $P$  is an element of  $X_0(p^2)(K)$  for some quadratic number field  $K$ ) then its  $j$ -height satisfies*

$$h_j(P) < C \cdot p^{13} \quad (100)$$

or even, under [4] (cf. Proposition 6.1),

$$h_j(P) < C \cdot p^{12} \log p. \quad (101)$$

**Proof** By Lemma 7.4, one can deduces from  $P$  a point  $P'$  in  $X_0(p)(K)$  which does not induce an element of  $X_0^+(p)(\mathbb{Q})$ , and whose  $j$ -height, say, is equal to  $h_j(P) + O(\log p)$  for an explicit function  $O(\log p)$  (see e.g. [51], inequality (51) on p. 240 and [6], Proposition 4.4 (i)). Replace  $P$  by  $P'$  if necessary. By Theorem 4.3, it is now sufficient to prove that  $h_\Theta(P - \infty) = O(p^{12} \log p)$  (focusing only on variant (101) of our height bounds).

Keep notations of Lemma 7.2. By construction, the point:

$$a := \varphi_e(P) + \varphi_e({}^\sigma P) = \varphi_e(P) - \varphi_e(w_p({}^\sigma P)) = \varphi_e(P - w_p({}^\sigma P))$$

is torsion. We first assume  $a = 0$ . Set  $X^{(2),-} := \{\iota_\infty(x) - \iota_\infty(y), (x, y) \in X_0(p)^2\}$  as in Proposition 5.2. Recall from Section 2 that  $\tilde{I}_{J_e^\perp, N_e^\perp}: J_e^\perp \rightarrow \tilde{J}_e^\perp$  is the map defined as in (1), that  $\iota_{\tilde{J}_e^\perp, N_e^\perp}$  is the embedding  $\tilde{J}_e^\perp \hookrightarrow J_0(p)$ , and denote by  $[N_{\tilde{J}_e^\perp}]_{\tilde{J}_e^\perp}$  the multiplication by  $N_{\tilde{J}_e^\perp}$  restricted to  $\tilde{J}_e^\perp$ . Define

$$\tilde{X}^{(2),-} := \left( \iota_{\tilde{J}_e^\perp, N_e^\perp} [N_{\tilde{J}_e^\perp}]_{\tilde{J}_e^\perp}^{-1} \tilde{I}_{J_e^\perp, N_e^\perp} \pi_{J_e^\perp} (X^{(2),-}) \right)$$

(which would just be the projection of  $X^{(2),-}$  on  $J_e^\perp$  if  $J_0(p)$  were *isomorphic* to the product  $J_e \times J_e^\perp$ , and is the best approximation to that projection in our case when the decomposition of the jacobian is true only up to isogeny). Then  $P - w_p({}^\sigma P)$  belongs to  $X^{(2),-} \cap \tilde{J}_e^\perp$ , and even to the intersection of surfaces (in the generic fiber):

$$X^{(2),-} \cap \tilde{X}^{(2),-}.$$

Note that  $\tilde{X}^{(2),-}$  is a priori highly non-connected, being the inverse image of multiplication by  $N_{\tilde{J}_e^\perp}$  in  $\tilde{J}_e^\perp$  of the (irreducible) surface  $\tilde{I}_{J_e^\perp, N_e^\perp} \pi_{J_e^\perp} (X^{(2),-})$ . However, in what follows we can replace  $\tilde{X}^{(2),-}$  by one of its connected components containing  $P - w_p({}^\sigma P)$ . Denote that component by  $\tilde{X}_P^{(2),-}$ .

By construction, the degree and height of  $\tilde{X}_P^{(2),-}$ , as an irreducible subvariety of  $J_0(p)$  endowed with  $h_\Theta$ , are those of  $\pi_{J_e^\perp} (X^{(2),-}) = X_{e^\perp}^{(2),-}$  relative to the only natural hermitian sheaf of  $J_e^\perp$ , that is, the  $\Theta_e^\perp = \Theta_{J_e^\perp}$  described in paragraph 2.1.2. One can therefore apply Proposition 5.2 to obtain



that all degrees are  $O(p^2)$ , all heights are  $O(p^2 \log p)$ . We claim the dimension of  $(X^{(2),-} \cap \widetilde{X}_P^{(2),-})$  is zero. That intersection indeed corresponds to couples of (different) points on  $X_0(p)$  having same image under  $\varphi_e$ . On the other hand, Brumer's conjecture implies  $X_0(p) \rightarrow \varphi_e(X_0(p))$  has generic degree one (cf. Lemma 7.2), so our intersection points correspond to singular points on  $\varphi_e(X_0(p))$ , which of course make a finite set.

We therefore are in position to apply our arithmetic Bézout theorem (Proposition 6.1), which yields  $h_\Theta(P - w_p(\sigma P)) \leq O(p^{12} \log p)$ . The two points  $(P - \infty)$  and  $(w_p(\sigma P) - \infty)$  have same  $\Theta$ -height (recall  $w_p$  is an isometry on  $J_0(p)$  for  $h_\Theta$ , cf. end of Remark 4.5), and of course are constructed to be different (Lemma 7.4). So one can apply them Mumford's repulsion principle (Proposition 5.7) to obtain

$$h_\Theta(P - \infty) \leq O(p^{12} \log p). \quad (102)$$

Let us finally deal with the case when the torsion point  $a = \varphi_e(P) + \varphi_e(\sigma P)$  is nonzero. We adapt the previous argument: pick a lift  $\tilde{a} \in J_0(p)(\overline{\mathbb{Q}})$  of  $a$  by  $\pi_e^\perp$  which also is torsion, and let  $t_{\tilde{a}}$  be the translation by  $\tilde{a}$  in  $J_0(p)$ . Replace  $(P - w_p(\sigma P))$  by  $t_{\tilde{a}}^*(P - w_p(\sigma P))$ ,  $X^{(2),-}$  by  $t_{\tilde{a}}^*X^{(2),-}$  and  $\widetilde{X}^{(2),-}$  by

$$\widetilde{t_{\tilde{a}}^*X}^{(2),-} = \iota_{\tilde{J}_e^\perp, N_e^\perp} [N_{\tilde{J}_e^\perp}]_{\tilde{J}_e^\perp}^{-1} \tilde{I}_{\tilde{J}_e^\perp, N_e^\perp} \pi_{J_e^\perp} (t_{\tilde{a}}^*X^{(2),-}).$$

Now  $t_{\tilde{a}}^*(P - w_p(\sigma P))$  belongs to  $(t_{\tilde{a}}^*X^{(2),-} \cap \widetilde{t_{\tilde{a}}^*X}^{(2),-})$ . The degree and height of  $t_{\tilde{a}}^*X^{(2),-}$  and  $\widetilde{t_{\tilde{a}}^*X}^{(2),-}$  (or rather, as above, some connected component  $\widetilde{t_{\tilde{a}}^*X}_P^{(2),-}$  of it containing  $t_{\tilde{a}}^*(P - w_p(\sigma P))$ ) are the same as for the former objects (in the case  $a = 0$ ). The fact that the intersection

$$t_{\tilde{a}}^*X^{(2),-} \cap \widetilde{t_{\tilde{a}}^*X}_P^{(2),-}$$

is zero-dimensional comes from the fact that otherwise, we would have  $\varphi_e(X_0(p)) = a - \varphi_e(X_0(p))$ , a contradiction with our present hypothesis  $a \neq 0$  by Proposition 7.2. The height bound for  $P$  is therefore the same as (102).  $\square$

**Corollary 7.6** *Under the assumptions of Theorem 7.5, if  $p$  is a large enough prime number and  $P$  is a quadratic point of  $X_0(p^\gamma)$  for some integer  $\gamma$  then  $\gamma \leq 25$ .*

**Proof** Combine Theorem 7.5 with isogeny bounds as in [22].  $\square$

**Remark 7.7** A similar (but technically simpler) approach for the morphism  $X_0(p) \rightarrow J_e$  over  $\mathbb{Q}$  should give (independently of any conjecture) a bound of shape  $O(p^7)$  for the  $j$ -height of  $\mathbb{Q}$ -rational (non-cuspidal) points of  $X_0(p)$  (which are known not to exist for  $p > 163$  by Mazur's theorem). The same should apply for  $\mathbb{Q}$ -points of  $X_{\text{split}}(p)$  (and here again, we obtain a weak version of known results).

## References

- [1] A. Abbes, Hauteurs et discrétude, Séminaire Bourbaki 1996-1997, exposé **825**, Astérisque **245** (1997), 141–166.
- [2] A. Abbes, E. Ullmo, Comparaison des métriques d'Arakelov et de Poincaré sur  $X_0(N)$ , Duke Math. J. **80** (1995), 295–307.
- [3] A. Aryasomayajula, Bounds for Green's functions on hyperbolic Riemann surfaces of finite volume. PhD thesis, Humboldt-Universität zu Berlin, 2013.
- [4] P. Autissier, An upper bound for the theta function, preprint (2015).
- [5] Yu. Bilu, P. Parent, Runge's method and modular curves, Int. Math. Res. Not. 2011, no. 9, 1997–2027.
- [6] Yu. Bilu, P. Parent, M. Rebolledo, Rational points on  $X_0^+(p^r)$ , Ann. Inst. Fourier **63**, n. 3 (2013), 957–984.
- [7] C. Birkenhake, H. Lange, Complex abelian varieties, Grund. der Mat. Wiss **302**, 2004.
- [8] J.-B. Bost, Intrinsic heights on stable varieties and abelian varieties, Duke math. J. **82** n. 1 (1996), 21–70.

- [9] *J.-B. Bost*, Potential theory and Lefschetz theorems for arithmetic surfaces. *Ann. Sci. École Norm. Sup.* (4), **32** (2) (1999), 241–312.
- [10] *J.-B. Bost, H. Gillet, C. Soulé*, Heights of projective varieties and positive Green forms. *J. Amer. Math. Soc.* **7** (1994), 903–1027.
- [11] *P. Bruin*, Explicit bounds on automorphic and canonical Green functions of Fuchsian groups, *Mathematika* **60** (2014), 257–306.
- [12] *A. Brumer*, The rank of  $J_0(N)$ , *Astérisque* **228** (1995), 41–68.
- [13] *S. Checcoli, F. Veneziano, E. Viada*, The explicit Mordell conjecture for families of curves, Preprint (2016). <https://arxiv.org/abs/1602.04097>
- [14] *T. Chinburg*, An introduction to Arakelov intersection theory, in *Arithmetic Geometry*, ed. by G. Cornell and J. Silverman, Springer-Verlag (1986).
- [15] *J.-M. Couveignes, B. Edixhoven et al.*, Computational aspects of modular forms and Galois representations, *Ann. of Math. Stud.* **176**, Princeton Univ. Press, Princeton, NJ, 2011.
- [16] *H. Darmon, L. Merel*, Winding quotients and some variants of Fermat’s last theorem. *J. Reine Angew. Math.* **490** (1997), 81–100.
- [17] *P. Deligne, M. Rapoport*, Les schémas de modules de courbes elliptiques, in “Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972)”, pp. 143–316; *Lecture Notes in Math.* **349**, Springer, Berlin, 1973.
- [18] *B. Edixhoven, R. de Jong*, Short introduction to heights and Arakelov theory, in *Computational aspects of modular forms and Galois representations*, *Ann. of Math. Stud.* **176**, Princeton Univ. Press, Princeton, NJ, 2011, pp. 79–94.
- [19] *B. Edixhoven, R. de Jong*, Applying Arakelov theory, in *Computational aspects of modular forms and Galois representations*, *Ann. of Math. Stud.* **176**, Princeton Univ. Press, Princeton, NJ, 2011, pp. 187–202.
- [20] *B. Edixhoven, R. de Jong*, Bounds for Arakelov invariants of modular curves, in *Computational aspects of modular forms and Galois representations*, *Ann. of Math. Stud.* **176**, Princeton Univ. Press, Princeton, NJ, 2011, pp. 217–256.
- [21] *G. Faltings*, Diophantine approximation on Abelian varieties, *Ann. of Math.* (2) **133** (1991), 549–576.
- [22] *É. Gaudron, G. Rémond*, Théorème des périodes et degrés minimaux d’isogénies, *Comment. Math. Helv.* **89** (2014), 343–403.
- [23] *É. Gaudron, G. Rémond*, Polarisation et isogénies, *Duke Math. J.*, **163** (2014), 2057–2108.
- [24] *Ph. Griffiths, J. Harris*, Principles of algebraic geometry. New York: Wiley Interscience, 1978.
- [25] *B. H. Gross*, Heegner points on  $X_0(N)$ , in *Modular forms* (ed. R.A. Rankin), Chichester: Ellis Horwood 87–106 (1984).
- [26] *B. H. Gross*, Heights and the special values of  $L$ -series. In *Number theory* (Montreal, Que., 1985), volume **7** of CMS Conf. Proc., pp. 115–187. Amer. Math. Soc., Providence, RI, 1987.
- [27] *B. H. Gross*, Heegner points and the modular curve of prime level, *J. Math. Soc. Japan*, **39** n. **2** (1987), 345–362.
- [28] *M. Hindry, J. H. Silverman*, Diophantine Geometry. An Introduction, GTM 201, Springer, 2000.
- [29] *H. Iwaniec, W. Luo, P. Sarnak*, Low lying zeros of families of  $L$ -functions, *Pub. Math. I.H.É.S.* **91** (2000), 55–131.
- [30] *H. Iwaniec, P. Sarnak*, The non-vanishing of central values of automorphic  $L$ -functions and Landau-Siegel zeros, *Israel J. Math.* **120** (2000), 155–177.
- [31] *J. Jorgenson, J. Kramer*, Bounds on canonical Green’s functions. *Compositio Math.* **142** (3) (2006), 679–700.
- [32] *M. A. Kenku, F. Momose*, Automorphism groups of the modular curves  $X_0(N)$ , *Composition math.* **65**, n. **1** (1988), 51–80.
- [33] *E. Kowalski, Ph. Michel, J. VanderKam*, Non-vanishing of high derivatives of automorphic  $L$ -functions at the center of the critical strip, *J. reine angew. Math.* **526** (2000), 1–34.
- [34] *E. Larson, D. Vaintrob*, Determinants of subquotients of Galois representations associated with abelian varieties. With an appendix by Brian Conrad. *J. Inst. Math. Jussieu* **13** (2014), no. 3, 517–559.
- [35] *S. Le Fourn*, Surjectivity of Galois representations associated with quadratic  $\mathbb{Q}$ -curves, *Math. Annalen* **365** (2016), 173–214. <http://arxiv.org/abs/1212.4713>
- [36] *Q. Liu*, Algebraic geometry and arithmetic curves. (Translated from the French by Reinie Ern ). Oxford University Press, Oxford, 2002.
- [37] *Yu. I. Manin, Yu. G. Zarhin*, Height on families of abelian varieties, *Mat. Sb. (N.S.)* **89** (131) (1972), 171–181.
- [38] *R. Menares*, Nombres d’intersection arithm tiques et op rateurs de Hecke sur les courbes modulaires  $X_0(N)$ , Th se de l’universit  de Paris-Sud-Orsay, 2008. <http://tel.archives-ouvertes.fr/tel-00360171>

- [39] *R. Menares*, Correspondences in Arakelov geometry and applications to the case of Hecke operators on modular curves, *Manuscripta Math.* **136** (2011), 501–543. arXiv:0911.0546
- [40] *F. Merkl*, An upper bound for Green functions on Riemann surfaces, in *Computational aspects of modular forms and Galois representations*, Ann. of Math. Stud. **176**, Princeton Univ. Press, Princeton, NJ, 2011, pp. 203–216.
- [41] *Ph. Michel, E. Ullmo*, Points de petite hauteur sur les courbes  $X_0(N)$ , *Invent. math.* **131** (1998), 645–674.
- [42] *F. Momose*, Isogenies of prime degree over number fields. *Compositio Math.*, **97** n. 3, (1995), 329–348.
- [43] *L. Moret-Bailly*, Métriques permises, in *Séminaire sur les pinceaux arithmétiques : la conjecture de Mordell*. *Astérisque* **127**, 29–87, 1985.
- [44] *L. Moret-Bailly*, Pinceaux de variétés abéliennes, *Astérisque* **129**, 1985.
- [45] *L. Moret-Bailly*, Sur l’équation fonctionnelle de la fonction thêta de Riemann, *Compositio math.* **75** n. 2 (1990), 203–217.
- [46] *D. Mumford*, On equations defining abelian varieties I, *Invent. Math.* **1** (1966), 287–354.
- [47] *D. Mumford*, Tata lectures on Theta Functions I, *Progr. in Math.*, No. **28**, Birkhäuser, 1984.
- [48] *A. Ogg*, Hyperelliptic modular curves, *Bull. Soc. Math. France* **102** (1974).
- [49] *A. Ogg*, Über die Automorphismengruppe von  $X_0(N)$ , *Math. Ann.* **228**, 279–292.
- [50] *F. Pazuki*, Theta height and Faltings height, *Bull. Soc. Math. Fr.* **140.1** (2012), 19–49.
- [51] *F. Pellarin*, Sur une majoration explicite pour un degré d’isogénie liant deux courbes elliptiques, *Acta Arith.* **100** (2001), 203–243.
- [52] *P. Philippon*, Sur des hauteurs alternatives, I [Math. Annalen **289** (1991), 255–283], II [Ann. Fourier, **44** (1994), 1043–1065], III. [J. Math. Pures Appl. **74** (1995), 345–365].
- [53] *G. Rémond*, Inégalité de Vojta en dimension supérieure, *Annali della Scuola Normale Superiore*, t. **29** n. 1 (2000), 101–151.
- [54] *G. Rémond*, Décompte dans une conjecture de Lang, *Invent. math.* **142** (2000), 513–545.
- [55] *G. Rémond*, Nombre de points rationnels des courbes, *Proc. London Math. Soc.* (3) **101** (2010), 759–794.
- [56] *K. Ribet*, Endomorphisms of semi-stable Abelian varieties over number fields, *Ann. of Math.* **101**, no. 3 (1975), pp. 555–562.
- [57] *E. Royer*, Petits zéros de fonctions  $L$  de formes modulaires, *Acta Arith.* **99** (2001), no 2, 147–172.
- [58] *Ch. Soulé*, Géométrie d’Arakelov et théorie des nombres transcendants, *Astérisque* **198–200** (1991), 355–371.
- [59] *L. Szpiro*, Un peu d’effectivité, in *Séminaire sur les pinceaux arithmétiques : la conjecture de Mordell*. *Astérisque* **127**, 275–287, 1985.
- [60] *E. Ullmo*, Hauteur de Faltings de quotients de  $J_0(N)$ , discriminants d’algèbres de Hecke et congruences entre formes modulaires, *American Journal of Math.* **122** (2000), 83–115.
- [61] *R. Wilms*, New explicit formulas for Faltings’ delta-invariant, Preprint (2016); arXiv:1605.00847v2.
- [62] *Sh.-W. Zhang*, Positive line bundles on arithmetic varieties, *J. Amer. Math. Soc.* **8** (1995), 187–221.

---

Pierre Parent  
IMB  
Université de Bordeaux  
351 cours de la libération  
33 405 Talence cedex France  
Pierre.Parent@math.u-bordeaux.fr